12

Technical Report 609

The Impact of CATTS System Characteristics on Selected Measures of Battalion Command Group Performance

Gary S. Thomas, Herbert F. Barber, and Ira T. Kaplan

ARI Field Unit at Fort Leavenworth, Kansas
Systems Research Laboratory

U. S. Army

Research Institute for the Behavioral and Social Sciences

January 1984

U. S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency under the Jurisdiction of the

Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Technical Director

L. NEALE COSBY
Colonel, IN
Commander

Technical review by:

Randall M. Chambers
James H. Banks
Edwin R. Smootz
Richard A. Kass

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: Technical Report 609	2. GOVT ACCESSION NO. AD-A140 332
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): THE IMPACT OF CATTS SYSTEM CHARACTERISTICS ON SELECTED MEASURES OF BATTALION COMMAND GROUP PERFORMANCE	5. TYPE OF REPORT & PERIOD COVERED: Final
	6. PERFORMING ORG. REPORT NUMBER: ARI FLv FU
7. AUTHOR(s): Gary S. Thomas, Herbert F. Barber, Ira T. Kaplan	8. CONTRACT OR GRANT NUMBER(s): --
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Army Research Institute for the Behavioral and Social Sciences, 5001 Eisenhower Avenue, Alexandria, VA 22333	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 2Q263744A795
11. CONTROLLING OFFICE NAME AND ADDRESS: Army Research Institute for the Behavioral and Social Sciences, 5001 Eisenhower Avenue, Alexandria, VA 22333	12. REPORT DATE: January 1984
	13. NUMBER OF PAGES: 72
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office): --	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE: --

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

--

18. SUPPLEMENTARY NOTES

A portion of this report appears in the Proceedings of the Human Factors Society, 27th Annual Meeting, Norfolk, VA, 1983.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Army Training & Evaluation Program (ARTEP)
Army Training Battle Simulation System (ARTBASS)
Battalion command group
Battle simulation
Battlefield performance
Command and control training
Information flow
Performance evaluation
Performance measurement

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This research was the first in a series of experiments recommended by ARI to support the ARTBASS Training Development Study (TDS). It investigated the effect of several system and scenario characteristics on battalion command group performance in a computer-driven battle simulation, CATTS. Results indicated that combat ratio, weather, mission type, and jamming of communication nets significantly affected simulated battle outcomes, but had no effect on ratings of command group behavior as assessed by a list of ARTEP subtasks. Ratings of command group performance made by CATTS controllers and (Continued)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Item 20 (Continued)

the players themselves increased with repeated exposure to CATTS training. A measure of intra- and inter-staff information flow indicated a small, non-significant improvement across exercise days. Results are discussed in terms of implications for future research supporting the ARTBASS TDS and training development in command and control procedures.

Technical Report 609

The Impact of CATTS System Characteristics on Selected Measures of Battalion Command Group Performance

Gary S. Thomas, Herbert F. Barber, and Ira T. Kaplan

Submitted by
Robert S. Andrews, Chief
ARI Field Unit at Fort Leavenworth, Kansas

Approved as technically adequate
and submitted for publication by
Jerrold M. Levine, Director
Systems Research Laboratory

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES
5001 Eisenhower Avenue, Alexandria, Virginia 22333

Office, Deputy Chief of Staff for Personnel
Department of the Army

January 1984

Army Project Number
2Q263744A795

Training and Simulation

ARI Research Reports and Technical Reports are intended for sponsors of R&D tasks and for other research and military agencies. Any findings ready for implementation at the time of publication are presented in the last part of the Brief. Upon completion of a major phase of the task, formal recommendations for official action normally are conveyed to appropriate military agencies by briefing or Disposition Form.

FOREWORD

The U.S. Army Research Institute, Fort Leavenworth Field Unit conducts a systems and training research program in support of the Combined Arms Center (CAC). The current experiment is one of several research projects completed that addressed command and control training and measurement in the computer-driven battle simulation, CATTS. This research was the first in a series of experiments recommended by the Field Unit to support the ARTBASS Training Development Study, and was designed to identify system and scenario characteristics that affect battalion command group performance. This report represents a more exhaustive treatment of the results that were presented in the form of a Field Unit Working Paper to the research sponsors, the ARTBASS proponents, for their immediate use.

EDGAR M. JOHNSON
Technical Director

THE IMPACT OF CATTS SYSTEM CHARACTERISTICS ON SELECTED MEASURES OF BATTALION COMMAND GROUP PERFORMANCE

EXECUTIVE SUMMARY

Objective:

The Combined Arms Center at Fort Leavenworth, Kansas has responsibility for the development of command and control (C^2) training systems. One of the most promising approaches to command and control training is the use of automation to develop computer-driven, C^2 training systems. Currently, the Combined Arms Tactical Training Simulator (CATTS) is the most advanced of these systems. Drawing on the information gathered from the operation of CATTS, the Army Training Battle Simulation System (ARTBASS), a mobile computerized C^2 training system, is now being developed. Prior to the fielding of ARTBASS, it will be necessary to demonstrate its training effectiveness. Due to the compressed developmental schedule of ARTBASS and a close similarity of ARTBASS to CATTS, the ARTBASS Test Integration Working Group decided that a Training Development Study (TDS) would be conducted using CATTS in lieu of ARTBASS.

In support of the CATTS/ARTBASS development and the TDS, the ARI Field Unit at Fort Leavenworth, has focused on refining measurements of command and control behaviors, the identification of variables that influence the difficulty level of CATTS/ARTBASS exercises, and on the measurement of changes in performance as command groups progress through CATTS training. The current project was the first in a series of three experiments recommended by ARI for the TDS. This project was designed to identify system and scenario characteristics which affect performance in simulated battle. The second project would address behavioral variables that influence battalion command group (BCG) effectiveness and identify appropriate training strategies to increase the training benefit of CATTS. The final project would attempt to maximize the training potential of CATTS exercises by combining the knowledge gained from the two previous projects to provide an optimum set of training exercises for battalion command groups.

The present research had the following objectives: (1) improved C^2 performance measurement, (2) determining the impact of various system and scenario characteristics on battalion command groups' perceptions and behaviors, and (3) obtaining command group performance data in support of the ARTBASS TDS. The results will assist future research and training efforts where specifications of exercise difficulty and measurement of BCG performance are required.

Procedures:

During the period October through December 1981, five battalion command groups each participated in four controlled one-day CATTS exercises. All five command groups were from mechanized infantry battalions and were matched as closely as possible in terms of size, previous experience, and composition of command group. Each exercise included a specific combination of five system and scenario characteristics selected for investigation. The system and scenario characteristics were weather (clear visibility vs. reduced visibility), terrain (Fulda vs. Sinai), communications (light vs. heavy jamming), mission (covering force vs. attack), and combat ratio (good vs. bad).

Exercises were driven by the CATTS computer, which moved units on the battlefield, calculated battle losses, reported unit status and casualties to company commanders, etc. Company commanders relayed information to the BCG staff and implemented BCG decisions on the battlefield through interactors, who interfaced directly with the computer. CATTS controllers role-played higher and adjacent staff members and generally supervised the training exercises.

During the investigation three types of measures were obtained concerning command group performance:

1. Ratings from the command groups themselves (players), controllers, and player-controllers (company commanders/FISTS). These raters also judged exercise difficulty and realism.

2. Simulation outcomes generated by the interaction of the battalion command groups with the computer through their company commanders and the CATTS interactors.

3. Communication was measured by a questionnaire designed to determine how much information command group members received and transmitted.

Findings:

Simulation outcomes and ratings of performance on a list of ARTEP subtasks were responsive to different aspects of CATTS training exercises. Changes in the battlefield environment significantly affected simulation outcomes, and controller and player self-ratings of BCG C^2 performance significantly increased as a function of repeated exposures to CATTS training. However, performance as assessed by ratings did not correlate with battlefield performance as measured by simulation outcomes. More specifically:

- Controller ratings and player self-ratings of performance on a list of ARTEP subtasks significantly increased across exercise days; whereas, player-controller ratings of BCG performance did not increase.

- Raters were unable to adequately discriminate among ARTEP subtasks as indicated by high inter-item correlations.

- Inter-rater agreement on performance ratings was confined to a subset of the controllers who rated BCGs.

- Rating groups did not agree in their ratings of exercise difficulty. Controllers perceived an increase across exercise day, players indicated no difference, and player-controllers indicated that the planning phase of exercises became progressively more difficult.

- All raters agreed that the realism of CATTS exercises was at least adequate and often very good.

- Good combat ratio, bad weather, reduced jamming of communications nets, and type of mission all contributed to better battlefield performance as measured by simulation outcomes.

- Intra- and inter-staff communication improved slightly but not significantly across exercise days.

Utilization of Findings:

The manipulations made in the system and scenario characteristics had a strong impact on battlefield performance as measured by simulation outcomes. These results have implications for the complexity of the battlefield environment simulated by CATTS, and the corresponding difficulty for BCGs operating in this environment. Not only is the CATTS battle calculus responsive to such manipulations, but these results demonstrated the need for setting levels of these variables so that different combinations of combat ratio, mission, weather, communications, and terrain can result in exercises of known levels of battle difficulty. Performance measurement research and training development efforts with subsequent BCGs can, thereby, be facilitated since performance change as a result of CATTS exercises can be attributed to training factors rather than the difficulty level attributed to simulated battle. In addition, it is in principle possible to set exercise difficulty to levels commensurate with the ability level of various BCGs to optimize trainings.

The detailed analyses of performance and difficulty ratings on the ARTEP questionnaire indicated that these subjective ratings were subject to the same problems of reliability and validity that are characteristic of most subjective ratings; therefore, these results should be interpreted with caution. Future research should be directed at developing better measures of C^2 process and product (battle outcome measures). Future performance assessment research should also focus on developing a link between battlefield performance as measured by simulation outcomes and the performance of C^2 processes by BCGs operating in a simulated battle environment. It is likely that performance on all the measures collected in this research could be enhanced by providing detailed feedback on the measures to BCGs operating in CATTS-like environments.

THE IMPACT OF CATTS SYSTEM CHARACTERISTICS
ON SELECTED MEASURES OF BATTALION COMMAND GROUP PERFORMANCE

CONTENTS

LIST OF TABLES

LIST OF FIGURES

Page

The Impact of CATTS System Characteristics On Selected Measures of Battalion Command Group Performance

INTRODUCTION

The advent of automation has generated new and expanded possibilities in the development of training vehicles for meeting the needs of the Army of the 1980's and 90's. One of the most promising uses of automation has been in the area of command and control (C^2) training. Specifically, the Combined Arms Tactical Training Simulator (CATTS), a computer driven free-play C^2 exercise system has been developed by the Combined Arms Center as a training vehicle and training development test bed. Drawing on information gathered during the development and refinement of CATTS, a follow-on system, the Army Training Battle Simulation System (ARTBASS), is now being developed.

Prior to fielding, it is necessary to demonstrate that ARTBASS is, in fact, effective for training battalion command groups in C^2 behaviors. The current development schedule for ARTBASS severely limits the amount of time available to determine its training effectiveness, but since ARTBASS in essence grew from CATTS, there is a close similarity between the two. Therefore, the ARTBASS Test Integration Working Group decided that a Training Development Study (TDS) could be conducted using CATTS in lieu of ARTBASS with the reasonable expectation that the effectiveness of ARTBASS would at least meet that of CATTS.

In support of CATTS/ARTBASS development and the TDS, the efforts of the ARI Field Unit at Fort Leavenworth included the refinement of C^2 measurement techniques, so that changes in C^2 performance as a function of exposure to CATTS could be documented. The current research was the first in a series of three experiments recommended by ARI to address the requirements of the TDS. It was designed to identify system and scenario variables which modify the difficulty of simulated battle in CATTS exercises; and therefore, the ability of BCGs to perform within this context. This was necessary because C^2 performance must be assessed relative to the context within which it is observed. For example, a given level of proficiency in executing C^2 procedures in a difficult and complex situation is superior to the level needed to succeed in an easy and simple training exercise. Identification of variables which impact on exercise difficulty is a first step in the development of C^2 training exercises at a known level of difficulty. This is a necessary element in the development of valid instruments and techniques to measure C^2 performance, so that performance changes due to CATTS training can be assessed.

A second experiment was proposed to identify behavioral variables that influence battalion command group (BCG) effectiveness and to identify training strategies that increase the value of CATTS exercises. This experiment would develop diagnostics to assess BCGs' competence in performing C^2 behaviors prior to CATTS exposure. It would then be possible to match exercise difficulty to the capabilities of BCGs coming to CATTS in a third experiment. This research would maximize the training potential of the CATTS

system by "tailoring" exercises to the ability level of BCGs and using known training enhancers such as detailed performance feedback. This final experiment was proposed to assess the potential effectiveness of CATTS/ARTBASS in training C^2 behaviors, and to maximize the training benefit of CATTS.

The primary focus of the present research was to develop measures of C^2 performance that are appropriate within the context of CATTS training exercises. Appropriate measures should be responsive to systematic manipulations in exercise difficulty and to training benefit derived by BCGs as a function of repeated exposures to CATTS training. To this end, the present research was to determine if, and to what degree, a selected set of simulated battle variables could affect the difficulty of CATTS exercises as assessed by battle outcomes, ratings of BCG performance, and ratings of exercise difficulty. Ratings of exercise realism were also collected from each group of raters as a check on the fidelity of the CATTS simulation.

The current research assessed the impact of system and scenario characteristics on aspects of BCG performance, so that these variables could be controlled or systematically varied in subsequent research. An attempt was made to minimize any systematic increase in BCG performance across exercise days since it would be a contaminating variable in the experiment. However, measures of C^2 performance were obtained that should be responsive to expected changes in BCG behaviors as a result of training. Although not the primary purpose of the project, C^2 performance data was gathered as a function of repeated exposure to CATTS, and may provide supporting evidence of training effectiveness for the ARTBASS TDS.

METHOD

Participants

During the period October through December 1981, five battalion command groups each participated in four controlled, one-day CATTS exercises. Attempts to match the battalion command groups and minimize command group differences as an extraneous variable were only partially successful due to the personnel turbulence in battalion command groups. All five command groups were from mechanized infantry battalions, with similar size and composition. Table 1 provides a list of the members of the command group (players) who were asked to participate in the exercises. Each key position (e.g., battalion commander, S1, S2, S3, and S4) was to be filled by the individual who normally occupied that position (i.e., incumbents). Several supporting members (not standardized) also participated in exercises.

Exercises

Each of the four exercises in which groups participated included a different combination of five system and scenario characteristics, so that no two exercises were identical. The system and scenario characteristics were terrain, mission, weather, communication, and combat ratio. Tables 2 and 3 describe in more detail the schedule of events and the various levels of each characteristic presented.

Table 1

Battalion Command Group Composition

Battalion Commander
S1
S4
S1 or S4 NCO
S1 or S4 RTO
S2
S2 NCO
S3
S3 Air
S3 NCO
S3 RTO
Company Commanders (1 tank, 2 line, 1 CSC)
Fire Support Officer
Fire Support NCO
Fist Chief (two)
Air Liaison Officer (Air Force)
Forward Air Controller (Air Force)

Table 2

Schedule of Events

Time	Day 1 Orientation	Day 2 Exercise 1	Day 3 Exercise 2	Day 4 Exercise 3	Day 5 Exercise 4
0730		Bde OPORD Planning	Bde OPORD Planning	Bde OPORD Planning	Bde OPORD Planning
1200		Bn OPORD	Bn OPORD	Bn OPORD	Bn OPORD
1300	Orientation	Execution	Execution	Execution	Execution
1400	Questionnaires				
1445	Player Controller Training				
1730	Release Players Controllers	Questionnaires *Feedback	Questionnaires *Feedback	Questionnaires *Feedback	Questionnaires *Feedback

*Feedback was limited to a discussion of tactics by company commanders for days 2 through 4. After the last exercise (day 5), an extensive feedback session was conducted.

Table 3

System and Scenario Characteristics

1.	Weather	Good - Relatively clear visibility or Bad - Reduced visibility due to excessive fog or dust
2.	Terrain	European - Fulda Gap or Middle East - Sinai Desert
3.	Communication	Good - Perfect commo, token jamming or Bad - Increased jamming, less landline, etc.
4.	Mission	Attack or Covering force operation
5.	Combat Ratio*	Good - Friendly 100% strength, enemy 75% strength or Bad - Friendly 80% strength, enemy 100% strength

*Percentages are approximations based on all assets available to enemy forces and maneuver assets available to friendly forces.

Experimental Design

The limited time and number of command groups available necessitated an efficient research design. The particular experimental design used in this study was a five-factor fractional factorial design (Daniel, 1976). As compared to a full factorial design, this design minimizes time and participant resources required to complete the experiment (exactly half that required by a full factorial design), but it sacrifices the number of degrees of freedom available. This design has been used in previous research to screen several variables simultaneously, so that the most influential can be investigated in further research. Use of the fractional factorial (F^2) design carries with it the acceptance of confounding between main effects and higher order interaction among the various independent variables (system and scenario characteristics). The design assumes that higher order interactions are not important. The specific combinations of variables presented to each command group is found in Table 4 along with the confounding between the main effects and higher order interactions (aliasing). The design also allows for the examination of two-way interactions involving one selected variable (in this case "mission" - variable D) since these interactions are confounded only with higher order interactions.

As indicated in the table, the design included a total of 16 combinations of variables (exactly half of the conditions required in a full factorial design), where four BCGs were exposed to four exercises. The fifth BCG was utilized as a "back-up" group in case of a computer malfunction causing the loss of simulation outcome data. This, in fact, occurred on exercise 1 for BCG II, so data from exercise 1 for BCG V was substituted for this mission data to complete the design. The data resulting from these 16 exercises was included in all analyses using simulation outcome measures, (see below). Analyses investigating ratings included all five BCGs and all 20 exercises. The information flow analysis was based on the first four BCGs.

Training System

Simulation. The battlefield environment was simulated by the CATTS computer, which provides a computer-driven exercise to train maneuver-battalion commanders and their staffs in the control and coordination of combined-arms operations. CATTS simulates the actions of units in combat, moves elements on and about the battlefield, calculates intervisibility and detection between forces, calculates weapon-to-target ranges, and the effects of weapons employment; it also maintains the status of personnel, equipment, ammunition, and fuel for friendly and enemy forces. Speed of movement, line of sight, and weapons effects are affected by changes in weather, terrain contour, soil type, suppressive fires, and personnel and equipment status. Given line of sight, engagements among maneuver weapon systems were automatic.

The CATTS exercise is conducted in a real-time, free-play mode. Within the prescribed tactical situation, the battalion commander can employ battalion assets in any manner deemed appropriate. The only constraints are the assets available to the battalion and the actions of the enemy commander. Deployment

of enemy assets were consistent across exercises within attack and covering force missions, but in accordance with threat doctrine, the threat controller made minor tactical adjustments to counter unique situations created by friendly force operations.

In this research, the command group, except the S1 and S4, occupied a simulated tactical operations center (TOC); the S1 and S4 were in another area, designated as the combat trains. The players (the battalion command group) in both areas were provided with communications equipment normally found in a maneuver battalion. They could communicate with higher, lower, and adjacent units (played by controllers) in any manner consistent with Army procedure and with the simulated location of the various units: face-to-face, by telephone, by radio, and by written message.

Figure 1 illustrates the communication among the players, the controllers, and the computer. Most communication took place by radio and telephone. The battalion command group had seven radio nets (actually hard-wired) with appropriate alternate frequencies. The nets included the following: the brigade command, brigade intelligence, brigade administration/logistics, battalion command, battalion administration/logistics, and air support nets. In addition, the command group also had a RATT (radioteletype) unit and field telephones, when appropriate. The sounds of enemy jamming, battle, and engine and generator noise were generated during the exercise to add to the realism of the experience.

Controllers. A team of controllers, permanently assigned to CATTS, mediated between the players and the computer. The control team consisted of a chief controller, who also played the role of brigade commander, S2 and S3, who played higher and adjacent positions, and S1 and S4 controllers who played the roles of higher and lower positions. In addition, a fire support controller, a direct air support controller, and a threat controller were also present. Two additional controllers identified as interactors, input orders into the computer at a control console: (a) the command and control interactor input orders from the battalion command group via company commanders to the maneuver units modeled in the computer; and (b) the fire support interactor input orders to the artillery and air support units. The threat actions were input directly by the threat controller, who, working independently, served as his own interactor.

Player-Controllers. Each command group brought along its company commanders and fire support representatives to serve as player-controllers. They received orders from battalion and translated them into subordinate unit maneuvers for input into the computer by the interactors. In addition, they also received battle status outputs from the computer and relayed that information back to the command group in the form of situation reports and spot reports.

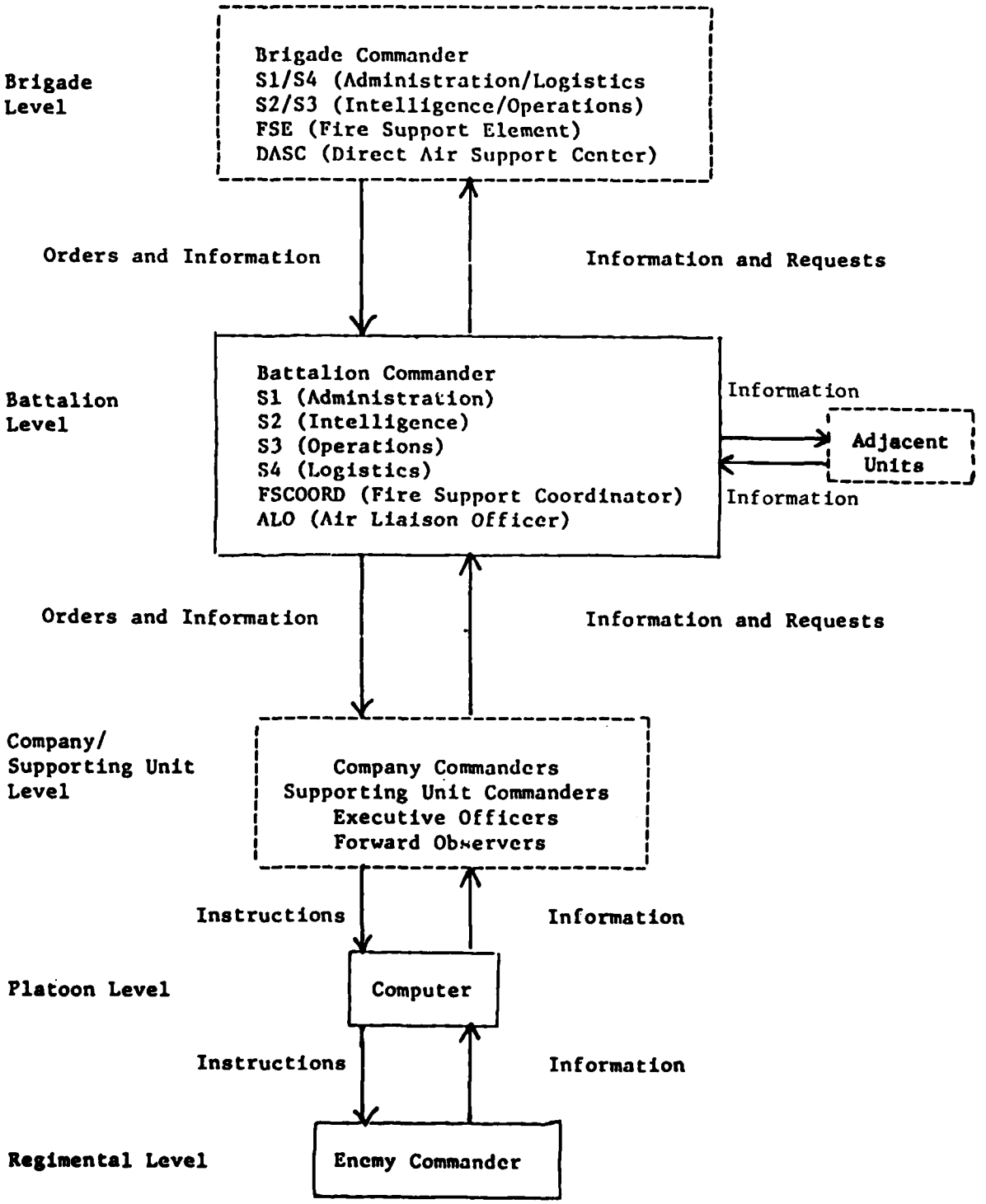

Figure 1. Communication between controller and player positions in CATTS. Controller positions are inclosed by broken lines.

Performance and System Measures

During this investigation, three measurement procedures were used: (1) ratings provided by the command group, player-controllers, and controllers; (2) simulation outcomes generated by the computer, and (3) information flow questionnaires designed to measure command group process.

Ratings. Ratings of battalion command group's performance and of exercise difficulty and realism during the CATTS exercises were obtained both from self estimates by the command group and from assessments by the player-controllers and controllers. Rating items were based upon a modified list of subtasks taken from the command group/staff module of ARTEP 71-2 (Army Training and Evaluation Program . . . 71-2, 1977). A list of the subtasks included in Appendix A was taken from previous ARI research (Barber and Kaplan, 1979; Kaplan and Barber, 1979; and Barber and Solick, 1980). The ratings were obtained using a magnitude estimation scaling technique. Magnitude estimation was used in an attempt to reduce scale compression and ceiling effects found with rating scales that used a limited number of discrete categories, where raters tend to use the upper range of the scale only. Magnitude estimation (Stevens, 1973) provides an open-ended scale so that a wider distribution of ratings can be obtained. Raters were asked to assess each subtask relative to a standard and assign a number to the subtask that reflected how many times greater or lesser it was than the standard. The standard, defined as the minimum acceptable performance in a tactical environment, was assigned the value of 100. All magnitude estimation raw scores were converted to $\log_{10}$ to normalize the distribution of scores prior to statistical analyses.

Performance estimates were obtained after each of the 20 CATTS training exercises from the controllers, players, and player-controllers. In addition, after each exercise, magnitude estimation ratings were obtained concerning the difficulty and realism of that exercise. These ratings of the functional areas of communication, intelligence, operations, administration/logistics, fire support, and coordination with other staff members along with an overall rating were obtained from players, player-controllers, and controllers for the planning phase and for the execution phase of the exercise. Sample rating forms are presented in Appendix A.

Battle simulation outcome measures. Modified versions of two objective indices obtained from combat development studies (USACDC, 1973) were used as measures of battlefield performance: Relative Exchange Ratio (RER) and Surviving Maneuver Force Ratio Differential (SMFRD). These two measures were selected because they were found to be significantly correlated with controller ratings of overall command group performance during previous CATTS exercises (Thomas, 1983). Two new measures were also significantly correlated with the controllers' ratings of BCG performance in that research: the change in combat ratio (ΔCR) and command and control index of lethality levels (C^2ILL). The latter measures were derived to provide indicators of overall performance in CATTS that generalize across mission and unit type. All four measures are mathematical relationships between OPFOR attrited or surviving and friendly forces lost or surviving. All data analysis involving these four simulation outcomes were conducted on the 16 exercises listed in Table 4. Table 5 presents the formulas for the four simulation outcome measures, and the rationale for using these measures appears in Appendix B.

Table 4

Five Factor Foldover Fractional Factorial Design for CATTS

Design:	A	B	C	D	E	
Unit	Weather	Terrain	Commo	Mission	Combat Ratio	Aliasing
Command Group 1						
Exercise 1	Good	Fulda	Good	Attack	Good	I=ABCE
Exercise 2	Good	Fulda	Bad	Covering Force	Bad	A=BCE
Exercise 3	Good	Sinai	Good	Covering Force	Bad	B=ACE
Exercise 4	Good	Sinai	Bad	Attack	Good	AB=CE
Command Group 2						
Exercise 1	Bad	Sinai	Bad	Attack	Bad	ABC=E
Exercise 2	Bad	Sinai	Good	Covering Force	Good	BC=AE
Exercise 3	Bad	Fulda	Bad	Covering Force	Good	AC=BE
Exercise 4	Bad	Fulda	Good	Attack	Bad	C=ABE
Command Group 3						
Exercise 1	Good	Fulda	Good	Covering Force	Good	D=ABCDE
Exercise 2	Good	Fulda	Bad	Attack	Bad	AD=BCDE
Exercise 3	Good	Sinai	Good	Attack	Bad	BD=ACDE
Exercise 4	Good	Sinai	Bad	Covering Force	Good	ABD=CDE
Command Group 4						
Exercise 1	Bad	Sinai	Bad	Covering Force	Bad	ABCD=DE
Exercise 2	Bad	Sinai	Good	Attack	Good	BCD=ADE
Exercise 3	Bad	Fulda	Bad	Attack	Good	ACD=BDE
Exercise 4	Bad	Fulda	Good	Covering Force	Bad	CD=ABDE

Table 5

Simulation Outcomes Computational Formulas

$$RER = \frac{\text{OPFOR Losses/OPFOR Initial Strength}}{\text{Friendly Losses/Friendly Initial Strength}}$$

$$SMFRD = \frac{\text{Friendly Remaining Strength}}{\text{Friendly Initial Strength}} - \frac{\text{OPFOR Remaining Strength}}{\text{OPFOR Initial Strength}}$$

$$\Delta CR = \frac{\dfrac{\text{OPFOR Initial Strength}}{\text{Friendly Initial Strength}} - \dfrac{\text{OPFOR Remaining Strength}}{\text{Friendly Remaining Strength}}}{\dfrac{\text{OPFOR Initial Strength}}{\text{Friendly Initial Strength}}}$$

$$C^2ILL = 1/2 \left(\frac{\text{Friendly Remaining Strength}}{\text{Friendly Initial Strength}}\right) + \left(\frac{\text{OPFOR Losses}}{\text{OPFOR Initial Strength}}\right)$$

Information flow. Measures of information flow were obtained by means of a questionnaire. The questionnaire was answered by six principal members of the command group: the battalion commander, S1 (administration), S2 (intelligence), S3 (operations), S4 (logistics), FSO (fire support officer). Four subordinate unit (company) commanders also completed this questionnaire. An example of this questionnaire is in Appendix C.

At the beginning of each exercise, the principal members of the battalion command group were briefed separately by their brigade counterparts, except for the commander and S3, who were briefed together. During these briefings, certain unique items of information were presented to each member. Then the members worked together for three to four hours to develop a plan, which they presented to their company commanders. Subsequently, the group members and company commanders answered a multiple-choice test based on the unique information originally presented in the brigade briefing. Their responses were analyzed to provide three measures of information flow:

1. Reception of required information that was presented to the respondent during the brigade briefing. The percentage of these items answered correctly was the individual's direct reception score.

2. Reception of required information that the group members should have received indirectly through those members who received it directly from brigade. The percentage of such items answered correctly was the intragroup communication score.

3. Reception of information required by the company commanders that should have been transmitted to them by members of the battalion command group. The percent of such items answered correctly measured communication to the company commanders.

Due to similarity of the various operations orders, it was not possible to obtain four different questionnaires. Therefore it was decided to measure information flow twice for each command group. Parallel forms of the questionnaire were administered on the second and fourth days for all five BCGs. While this procedure did not permit an examination of the effect of system and scenario characteristics, it could provide some, albeit incomplete, indication of improvement in information transfer across exercises.

Table 6 presents all the measures collected during the CATTS exercises.

RESULTS AND DISCUSSION

Preliminary Analysis of C^2 Performance Measures

A major objective of the current research was to examine the appropriateness of the various measures of performance within the context of the CATTS battle simulation. Three general categories of measurement instruments were used during this research effort: simulation outcomes; subjective ratings by players, controllers, and player-controllers; and a measure of information

Table 6

Data Collection Plan

	Day 1 Orientation	Day 2 Exercise 1	Day 3 Exercise 2	Day 4 Exercise 3	Day 5 Exercise 4
Player Ratings		X	X	X	X
Player-controller Ratings		X	X	X	X
Controller Ratings		X	X	X	X
Simulation Outcomes*		X	X	X	X
Information Flow Questionnaires		X		X	

* Simulation outcomes were recorded for the first 16 exercises where no computer failures occurred.

flow using a questionnaire. Kaplan (1980) examined the information flow questionnaire and its applicability to measurement of inter and intra-staff communication during the planning phase of CATTS exercises. The questionnaire remains basically unchanged from that earlier work, and the present research was to determine if such communication improves as a result of repeated exposure to CATTS training. Thomas (1983) performed an initial investigation of the validity of the simulation outcomes as measures of command group performance. Since the reliability and validity of these two measures have been previously examined, this preliminary analysis concentrates on the subjective ratings made by players, controllers, and player-controllers. Specifically, the ratings were analyzed to determine if the raters discriminated among items, and if the raters agreed on their assessment of battalion command group performance. Then, the relationship between ratings and simulation outcomes was examined. Following this preliminary analysis, the degree to which the various measures were responsive to systematic manipulations of CATTS system characteristics or responsive to expected C^2 training from CATTS is presented.

Analysis of Ratings on ARTEP Subtasks. Controllers, players, and player-controller ratings of performance and difficulty were analyzed to identify relationships in these data. First, inter-item correlations were calculated among the 27 items on the performance portion of the ARTEP questionnaire on all 20 exercises (see Appendix A). All inter-item correlations among the performance items (see Table 7) were significant ($p<.005$). The interitem correlations ranged from .33 to .98 with a median of .88 for controllers. Similarly, inter-item correlations for players ranged from .35 to .95 and player-controllers ranged from .37 to .92. However, the magnitude of the median correlations were not as high as in the case of the controllers (.54 and .62 respectively).

Similar analyses were performed on the exercise difficulty items, and again, nigh and significant inter-item correlations were observed for controllers, players, and player-controllers ratings. The median correlations were .85, .51, and .61 respectively, as indicated in Table 7.

The correlations between similar items on the performance and difficulty sections of the questionnaire were also calculated. Although the magnitude of these correlations were somewhat less than the performance inter-item correlations (medians = .67, .34, and .54 respectively), all the correlations were significant beyond the .005 level.

The high inter-item correlations among the performance and difficulty items, suggest there are some basic underlying dimensions or commonality across the various items. Therefore, the performance and difficulty ratings were subjected to factor analysis to determine whether some underlying variables (factors) could account for the observed covariation among the ratings. The number of factors derived varied with the rater type (see Appendix D). Only one factor was derived from performance ratings obtained from controllers, suggesting that they were unable to discriminate among the ratings on the various ARTEP subtasks. The factor analysis of the players ratings of performance resulted in a three factor solution, roughly equivelant to the task groupings of the ARTEP subtasks. Factor I was equivalent to Task

Table 7

Inter-item Correlations for Performance and Difficulty Items

		Among all 27 Performance Items	Among all 14 Difficulty Items	Between Corresponding Performance and Difficulty Items
Controllers	Median	.88	.85	.67
	Range	.33 - .98	.53 - .98	.43 - .89
	N*	41 - 124	85 - 128	78 - 126
Players	Median	.54	.51	.34
	Range	.35 - .95	.27 - .83	.11 - .54
	N*	144 - 191	117 - 159	117 - 158
Player-Controllers	Median	.62	.61	.54
	Range	.37 - .92	.35 - .87	.36 - .74
	N*	101 - 139	78 - 115	76 - 107

* Difference in N-sizes were due to the fact that not all raters rated all items.

I: "Gather and analyze required information." Factor II was nearly equivalent to the combined tasks 2 and 3: "Develop a plan based on mission and modify it as required by events," and "Communicate/coordinate." The last factor was roughly equivalent to task 4 and 5: "Implement plan," and "Supervise combat operations." The factor analysis of the performance ratings recorded by player-controllers resulted in two factors: Factor I representing task 4 and 5 and 4 subtask items from task 2; and Factor II equivalent to tasks 1 and 3 plus 8 subtasks of task 2.

The factor analysis performed on exercise difficulty questions indicated planning difficulty and execution difficulty factors for both the controllers and player-controllers. That is, Factor I included all items referring to planning, and Factor II included all items assessing the difficulty of execution. The analysis of player ratings of exercise difficulty also resulted in two factors. However, Factor I included all items except those assessed in the difficulty of planning and execution of commo and admin/log.

The small number of factors and the relatively high inter-item correlations suggest that the raters had difficulty discriminating among the ARTEP items intended to assess battalion command group performance and exercise difficulty. However, it appears that the players and player-controllers are better able to make these discriminations.

Inter-rater Agreement. In addition to analyzing the relationship between items, the extent to which raters agreed among themselves in the performance of command groups was also examined. First, inter-rater agreement was calculated by comparing relative ratings on the 27 ARTEP performance items for each exercise day both among and between rater groups. Correlation were highly variable (ranged from -.97 to +.94) and on average, low (median = .10). There were no consistent patterns among any subset of raters. These findings are consistent with the earlier results suggesting that the raters could not make the fine item discriminations required by the instruments.

The lack of consistent agreement on the individual ARTEP items suggested an examination of inter-rater agreement on the composite factor scores. However, while factor scores may be the appropriate level of analysis, only controllers observed enough exercises to allow for meaningful agreement comparisons. In the case of controllers, one factor score (i.e., a mean score taken across all 27 performance items) was calculated for each exercise day for each controller. Inter-rater agreement among controllers was assessed on this factor across all 20 exercises observed. Correlations ranged from -.30 to .96 with the median inter-rater agreement equal to .36. There was a tendancy for the brigade S2, S3, OPFOR and Chief controllers to agree among themselves as to the performance of the battalion command group as assessed by this factor rating. Their inter-rater correlations ranged from .66 to .96 with a median of .73. The brigade S1, S4, maneuver controller, and the Fire Direction Center Controller did not significantly or consistently agree among themselves or with any other controllers as to the performance of the battalion command groups.

While there was some general agreement within a subset of controllers, the fact remains that there was not a consistent agreement among all controllers as to battalion command group performance. One probable explanation is that the controllers are involved in functions during the CATTS exercise that competed for time required to observe BCGs, and observed different aspects of the command group's performance. The general lack of inter-rater agreement among controllers (and by extrapolation players, and player-controllers) may also be due to differing interpretations of the ARTEP standards.

These findings coupled with the lack of item discriminability highlight the need for modification in the performance rating instruments. Only when an average was taken across all items within a factor, did there appear to be inter-rater agreement. Given the high commonality of the rated items and the improved inter-rater agreement, composite factor scores appeared to be more appropriate measures of performance and difficulty rather than individual items. Therefore, factor scores were used in the remainder of the report to examine and describe further trends in the data.

Comparisons Among Measures. The simulation outcome measures are all composites of percentages of OPFOR attrited or surviving and friendly forces lost or surviving, so relationships among these measures is expected. As indicated in Table 8, C^2ILL and SMFRD correlate significantly with all other combat measures. These results concur with the findings of Thomas (1983), who also reported high intermeasure correlations. However, that research also reported very high correlations between the simulation outcomes and controller ratings of BCG performance. Mean performance scores were calculated by converting all controller ratings of the ARTEP subtasks to Z-scores for each rater (to control for rater response bias) on each exercise day, and then summing across ratings for all controllers observing each exercise day. Previous findings were not replicated in the current research. As indicated in Table 8, the mean ratings of performance score did not correlate significantly with any of the simulation outcomes.

Since BCG performance must be viewed within the context of exercise difficulty (for example, a given level of performance in an easy exercise should not be considered as desirable as the same level of performance in a difficult exercise), the simulation outcomes could also in part be measures of exercise difficulty. Therefore, correlations between controllers ratings of exercise difficulty and the simulation outcomes were calculated and are also presented in Table 8. The Difficulty Planning and Difficulty Execution variables were calculated in the same manner as the mean performance score. The mean of Z-scores was calculated for each rater across items referring to planning and execution, respectively, and sums were taken across raters per exercise day. As shown in the table, there were no relationships between the simulation outcomes and controller ratings of exercise difficulty. It is possible that in the current research, these measures assessed different aspects of performance. There were, however, highly significant correlations between controller ratings of exercise difficulty and of ratings of BCG performance, suggesting a general bias in these ratings.

Table 8

Intercorrelations Among Simulation Outcomes and Controller Ratings on ARTEP Subtasks

	RER	SMFRD	C^2ILL	ΔCR	Factor I PERF	DIFF I	DIFF II	DAY
RER	-	.62*	.94**	.35	-.27	-.15	-.08	-.10
SMFRD		-	.81**	.93**	-.06	-.24	-.21	.07
C^2ILL				.60*	-.24	-.19	-.17	-.06
ΔCR				-	-.04	-.30	-.26	.08
Factor I Performance					-	.68*	.80**	.77**
Factor I Difficulty							.69*	.82**
Factor II Difficulty							-	.66*
Exercise Day								-

*p < .01
**p < .001

System Characteristics

The following discussion considers the impact of system characteristics on simulation outcomes and controller ratings of BCG performance and of exercise difficulty. Similar analyses were not conducted on player and player-controller ratings, since these raters did not observe all experiences; and therefore, raters and experimental conditions were confounded.

Simulation Outcomes. As stated previously the fractional factorial (F^2) design is intended to economically "screen" several variables concurrently to determine if the impact of main effects and selected interactions are of potential importance. The analysis allows one to determine the responsiveness of simulation outcomes to manipulations in independent variables (system characteristics). However, several two and three-way interactions are confounded; and therefore, are uninterpretable (see Table 4).

The mean values for each simulation outcome as a result of manipulations in system characteristics are presented in Table 9. To test for important differences in these means and for potential interactions involving the mission variable (mission type was considered the variable of most interest) seperate F^2 analyses were conducted for each simulation outcome. ANOVA tables for these analyses appear in Appendix E. Sums of squares of interactions that resulted in F-values of 1.0 or less were included in mean square error terms for the analyses, accounting for the differences in degrees of freedom for the error terms among the four analyses summarized in Table 10. This table includes the potentially important effects as defined by the fact that at least 5% of the variance in the analysis was accounted for by the effect.

The main effects and interactions listed for the RER measure account for 84% of the variance in the experiment, leaving only 16% accounted for by the other 10 effects combined. As shown in the table, all effects were significant ($P<.05$) or marginally significant ($P<.10$), even though degrees of freedom were relatively low. RER scores were higher in the covering force mission (where the defender had a tactical advantage) with favorable combat ratio, and in poor weather. There was a tendency for the Sinai covering force to be performed better than the Fulda covering force, and for the Sinai attack to be performed poorer than the Fulda attack. Finally, poor commo (or high jamming) had a greater effect in the covering force than in attack missions. This may be due to the fact that most units were on "radio listening silence" in the attack until the enemy was encountered.

The SMFRD measure was significantly affected by combat ratio and weather, where good combat ratio and bad weather resulted in better performance on the battlefield. As in the case of RER, mission and the mission-by-terrain interaction were important, each accounting for about 5% of the variance though not reaching statistical significance. However, scores were higher in the attack than in the covering force. The commo manipulation was again important, but in this situation, as a main effect.

Table 9

Mean Values For Simulation Outcomes
as a Function of Levels of System Characteristics

	Combat Ratio		Mission		Weather		Communication		Terrain	
	Good	Bad	Covering Force	Attack	Good	Bad	Good	Bad	Fulda	Sinai
RER	.43	.24	.43	.23	.28	.39	.36	.31	.33	.34
SMFRD	-.18	-.34	-.28	-.24	-.30	-.22	-.23	-.29	-.27	-.25
ΔCR	-.29	-.68	-.61	-.36	-.40	-.57	-.43	-.54	-.52	-.46
C^2ILL	.48	.39	.46	.41	.41	.46	.45	.42	.43	.43

Also indicated in Table 10, is the fact that the battlefield variables of combat ratio, mission, weather, commo, and the mission-by-terrain interactions had similar effects on the battlefield performance measure, C^2ILL, where covering force scores exceeded that in the attack. Finally, a smaller number of effects were important with regard to the ΔCR measure. Good combat ratio, bad weather, and attack resulted in higher performance scores. The mission-by-combat ratio interaction was significant indicating that bad combat ratio had a more detrimental effect in the covering force operation than in the attack mission.

The effects of the system variables on the measures of battlefield performance are relatively strong and consistent, where combat ratio, weather, mission type, commo, and the mission-by-terrain interactions were important effects. However, two of the simulation outcome measures indicate better performance in the attack, while two favor the covering force. These differences are probably artifactual due to the assumptions made in calculating initial enemy strength, which is a component in all formulas. In this research, all enemy resources within the friendly forces' sector of responsibility were considered part of the initial enemy strength even if they were not engaged in direct combat in attacks or covering force operations. Although there are some relative differences depending upon the battlefield measure under consideration, decreasing the amount of force considered at the disposal of the enemy prior to battle has the effect of increasing covering force scores with respect to attack scores to the point where the two distributions no longer overlap. At this point, of course, covering force performance would always exceed attack performance regardless of other factors.

Performance and Difficulty Ratings. Controller ratings of BCG performance and exercise difficulty were subjected to F^2 analyses to determine if these ratings were responsive to the manipulation of system variables. The mean performance ratings for each exercise day were calculated by averaging across all performance items for each controller, and then calculating a grand mean for each day. The same procedure was used to calculate execution and planning difficulty means. It was found that in none of the analyses were ratings affected by the system variables of combat ratio, mission type, weather, commo, or terrain. Means are presented in Appendix F.

Changes Across Exercise Days

Ratings. Factor scores were calculated for each exercise day by rater type (see Appendix G). Even though the ratings collected from controllers had limited inter-rater agreement, controller ratings did increase across exercise days as indicated in Figure 2. The mean performance ratings taken across all items for each exercise increased significantly across the four exercise days ($F_{3,16} = 9.051$, $P < .001$). A Tukey's HSD indicated that performance on the third and fourth exercise days was greater than performance on the first day ($P < .05$).

Table 10

Summary Statistics From the Fractional Factorial Analysis

RER				SMFRD				ΔCR				C^2ILL			
VAR	$F_{1,8}$	P	ETA^2	VAR	$F_{1,9}$	P	ETA^2	VAR	$F_{1,9}$	P	ETA^2	VAR	$F_{1,7}$	P	ETA^2
Combat Ratio	16.22	.005	.28	Combat Ratio	22.55	.005	.51	Combat Ratio	22.43	.005	.44	Combat Ratio	21.78	.005	.43
Mission	18.43	.005	.32	Mission	2.10	NS	.05	Mission	9.21	.025	.18	Mission	6.33	.05	.13
Weather	5.68	.05	.10	Weather	5.33	.05	.12	Weather	4.26	.10	.08	Weather	6.07	.05	.12
				Commo	2.55	NS	.06					Commo	3.82	.10	.08
Miss X Terrain	3.91	.10	.07	Miss X Terrain	2.35	NS	.05	Miss X Combat Ratio	3.89	.10	.08	Miss X Terrain	3.19	NS	.06
Miss X Commo	3.84	.10	.07												
Residual			.12	Residual			.15	Residual			.17	Residual			.13
Total variance accounted for by the above factors was .84.				Total variance accounted for by the above factors was .79.				Total variance accounted for by the above factors was .78.				Total variance accounted for by the above factors was .82.			

Controller ratings of exercise difficulty were analyzed for planning and execution related items, that are presented in Figure 3 under Factors I and II, respectively. The controllers reported an increase in the difficulty of execution ($F_{3,16} = 4.428$, $P<.05$) and planning ($F_{3,16} = 10.090$, $P< .001$) across exercise days. It is difficult to interpret these results, since no manipulations were made through the course of this experiment that would affect the difficulty of the planning phase, and the presentation order of system variables was counterbalanced to prevent a difference in the execution phase. It is possible that controller ratings were contaminated by a systematic bias to increase ratings across exercise days. It will be recalled that ratings of BCG performance and exercise difficulty correlated significantly with exercise days (Table 8). Since controllers were aware of the purpose of the TDS, it is possible that their ratings were influenced toward demonstrating training effectiveness of CATTS.

Player ratings of performance and difficulty were analyzed in the same fashion to determine if they perceived a change as a function of CATTS exposure. Since player self-ratings of performance grouped into three factors, these factors were analyzed separately (see Figure 4). Players indicated that they improved on all three factors ($F_{3,12} = 7.34$, $P< .005$, $F_{3,12} = 14.96$, $P<.001$; and $F_{3,12} = 4.06$, $P< .05$; respectively). Post-hoc tests indicated that performance on the third and fourth exercise days was significantly better than on the first exercise day ($P<.05$) for Factors I, and II. Performance on exercise day three was rated better than day one ($P<.05$) for Factor III. Players did not, however, perceive a change in difficulty across exercise days (Figure 5) on the two-factor categorization of their responses ($F_{3,12} = 1.80$, and $F_{3,12} = 1.70$). These findings taken together, support the assertation that players did indeed perceive an increase in performance as a function of repeated exposure to CATTS.

Player-controller ratings were also analyzed according to the factors derived from previous analyses. The performance ratings on Factors I and II (Figure 6) indicate no significant increase across exercise days ($F_{3,12} = 3.17$ and $F_{3,12} = 2.69$ respectively). As shown in Figure 7, player-controllers also indicated an increase in the difficulty of planning ($F_{3,12} = 3.57$, $P< .05$), but not in the difficulty of execution ($F_{3,12} = .50$). In general, player-controllers did not perceive a significant increase in BCG performance, but as in the case of controllers, player-controllers evidently had difficulty in discriminating among the ARTEP subtasks, and their ratings of performance and exercise difficulty were highly correlated (see Table 7). It is not clear why they observed an increase in planning difficulty; possibly they became more aware of the efforts involved in planning with repeated exposure to exercises.

Simulation Outcomes. There was no indication that BCG performance, as measured by the battle outcomes, increased with repeated exposure to CATTS (see Figure 8). The 2X2 ANOVA for each simulation outcome compared first and second exposure to attack and covering force missions. There was no tendency for second exposure scores to be higher than first exposure scores. There was a slight tendency, however, for attack scores to increase and covering force scores to decrease across exercises, but the interactions were not significant. Mean values appear in Appendix H.

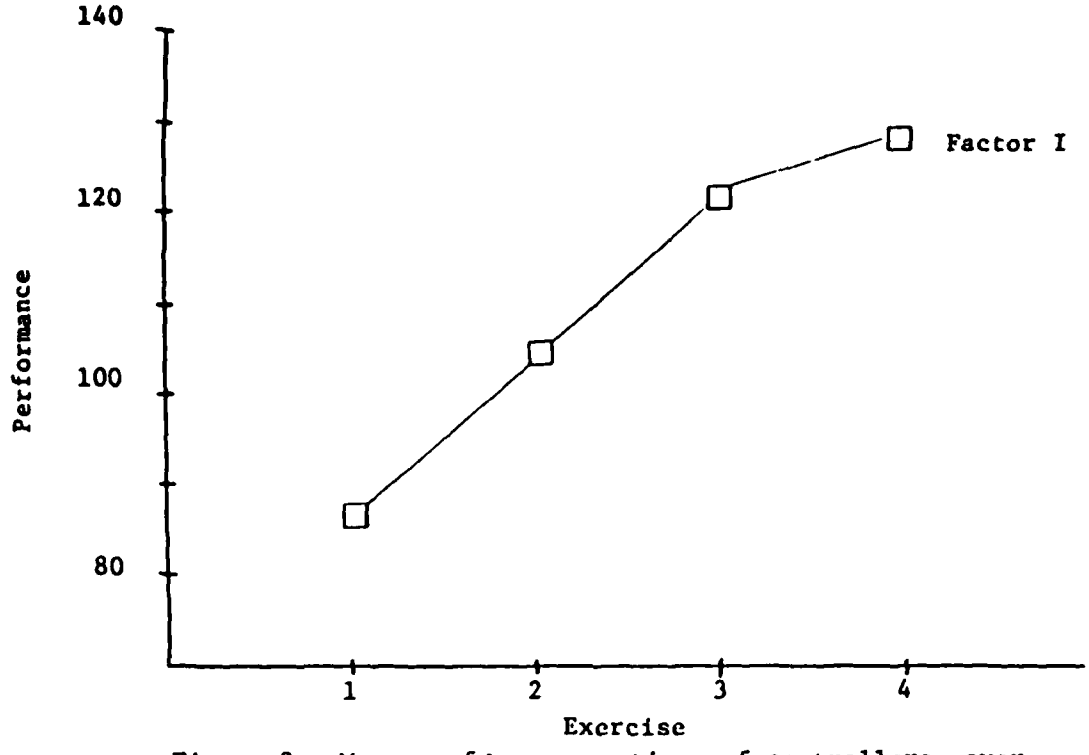

Figure 2. Mean performance ratings of controllers over exercise days.

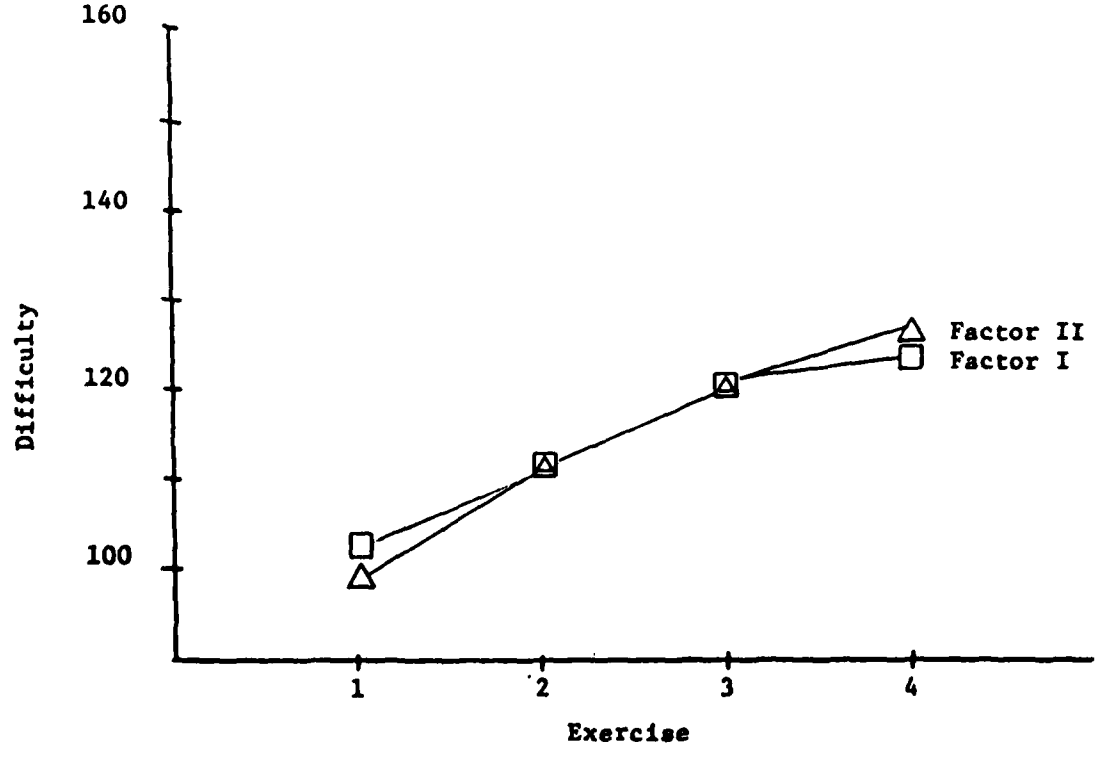

Figure 3. Mean difficulty ratings of controllers over exercise days.

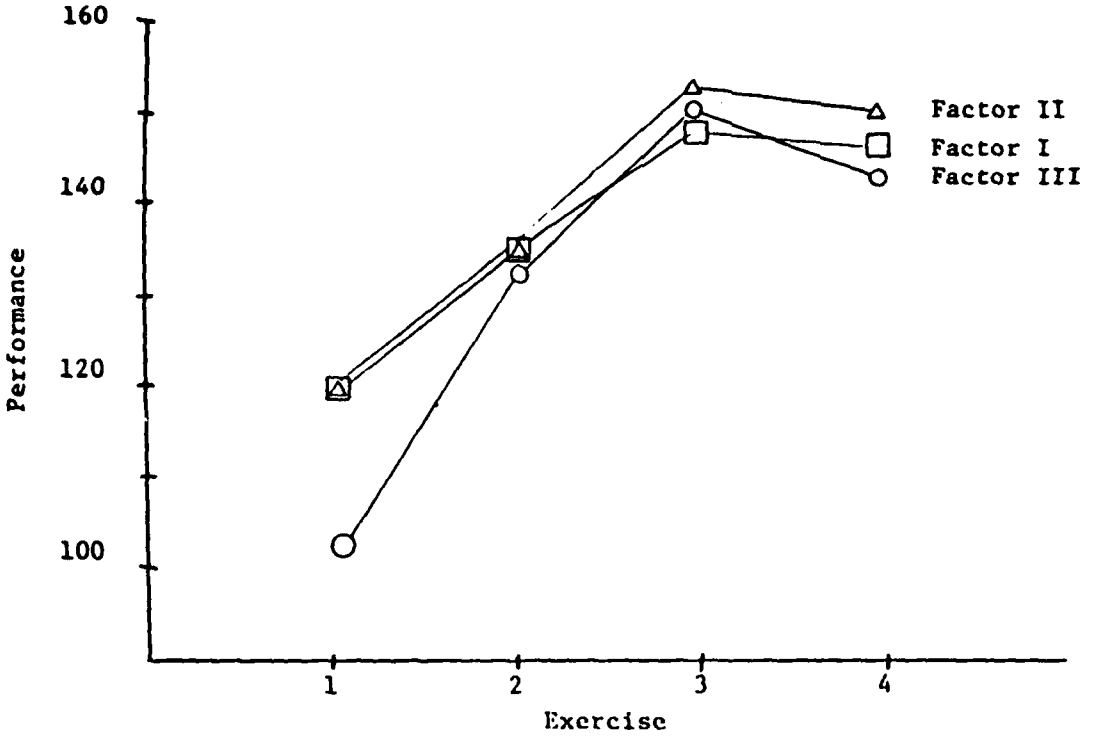

Figure 4. Mean performance ratings of players over exercise days.

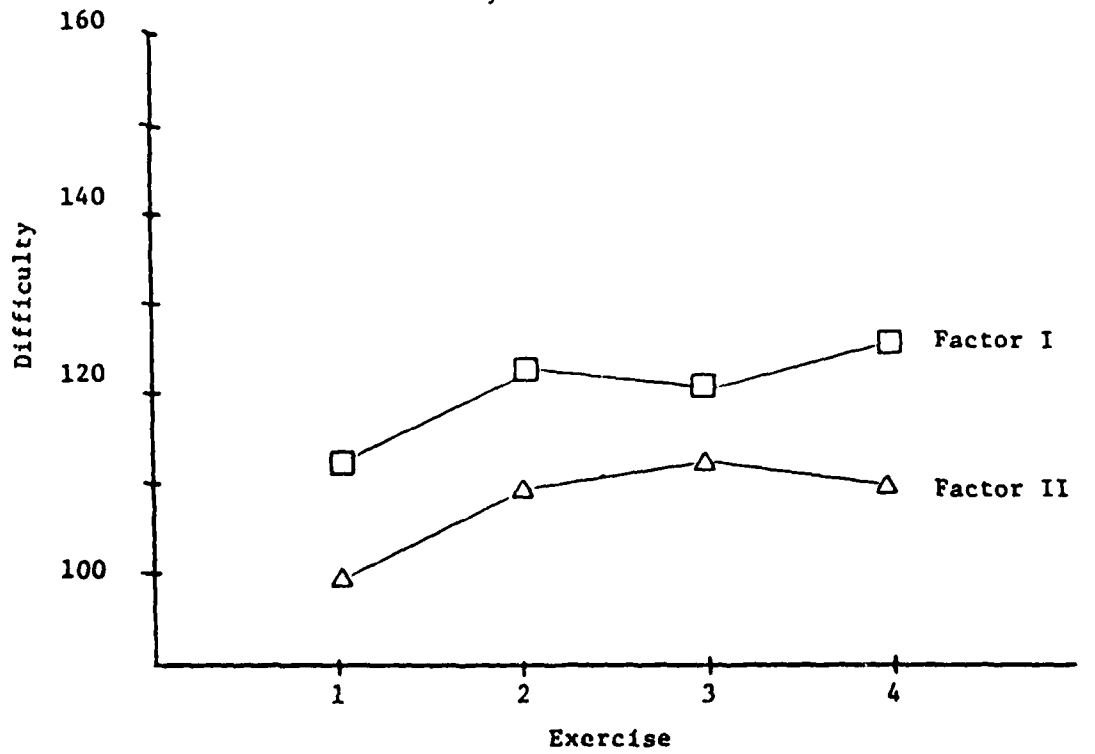

Figure 5. Mean difficulty ratings of players over exercise days.

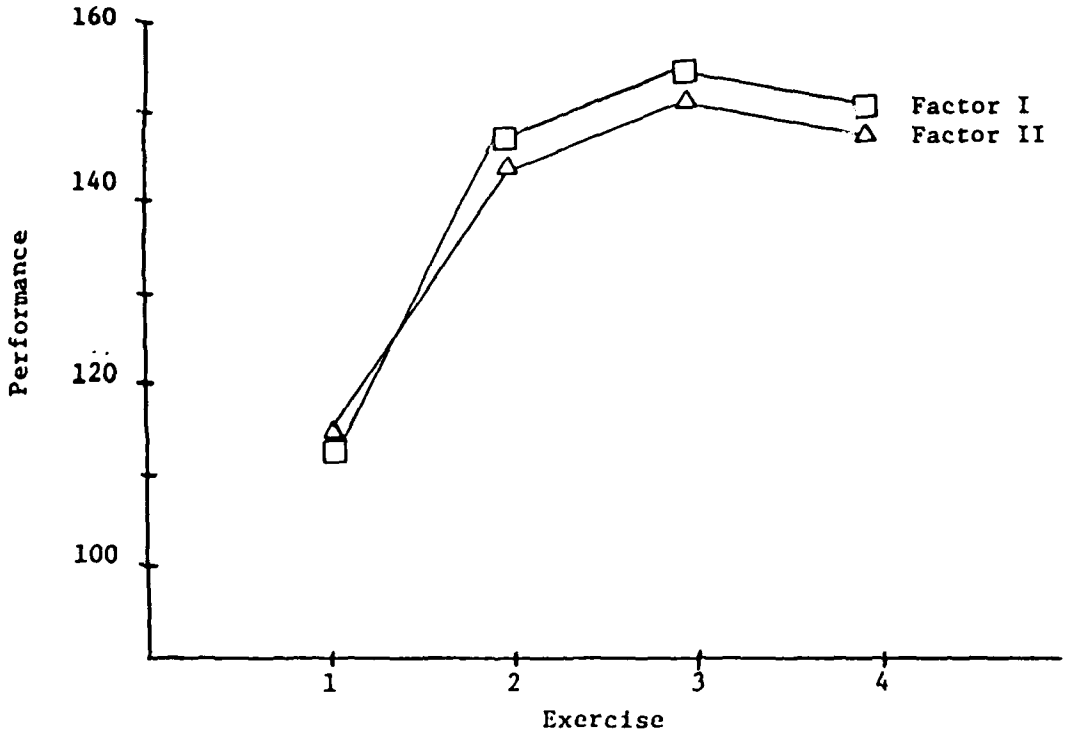

Figure 6. Mean performance ratings of player-controllers over exercise days.

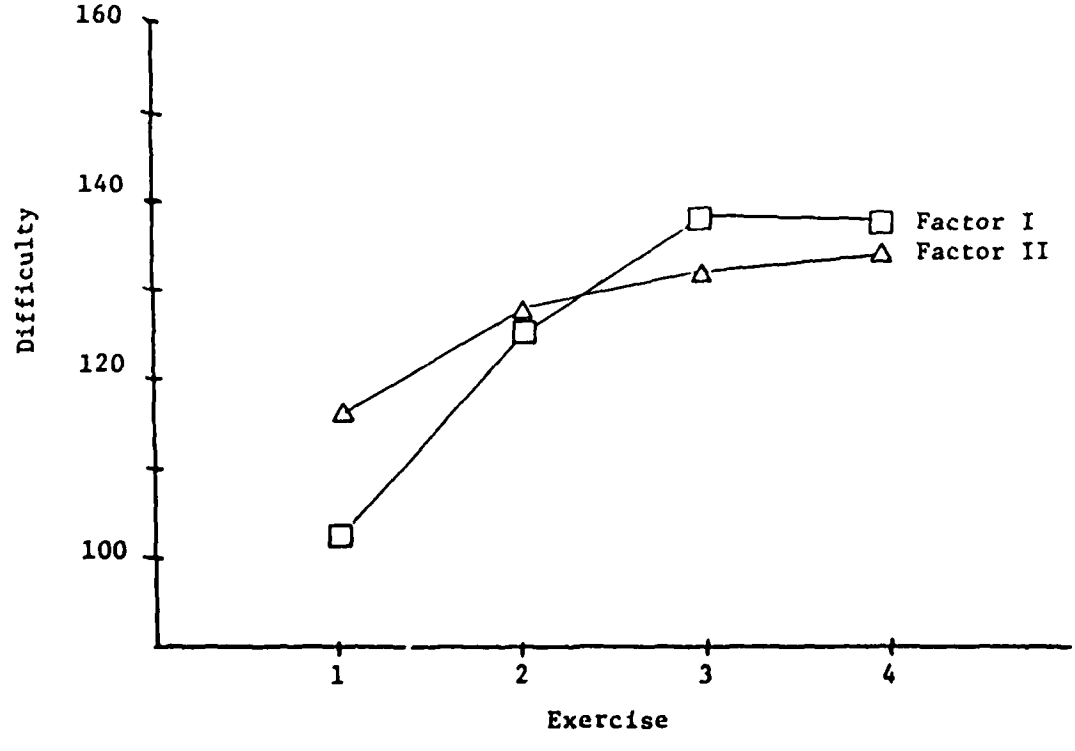

Figure 7. Mean difficulty ratings of player-controllers over exercise days.

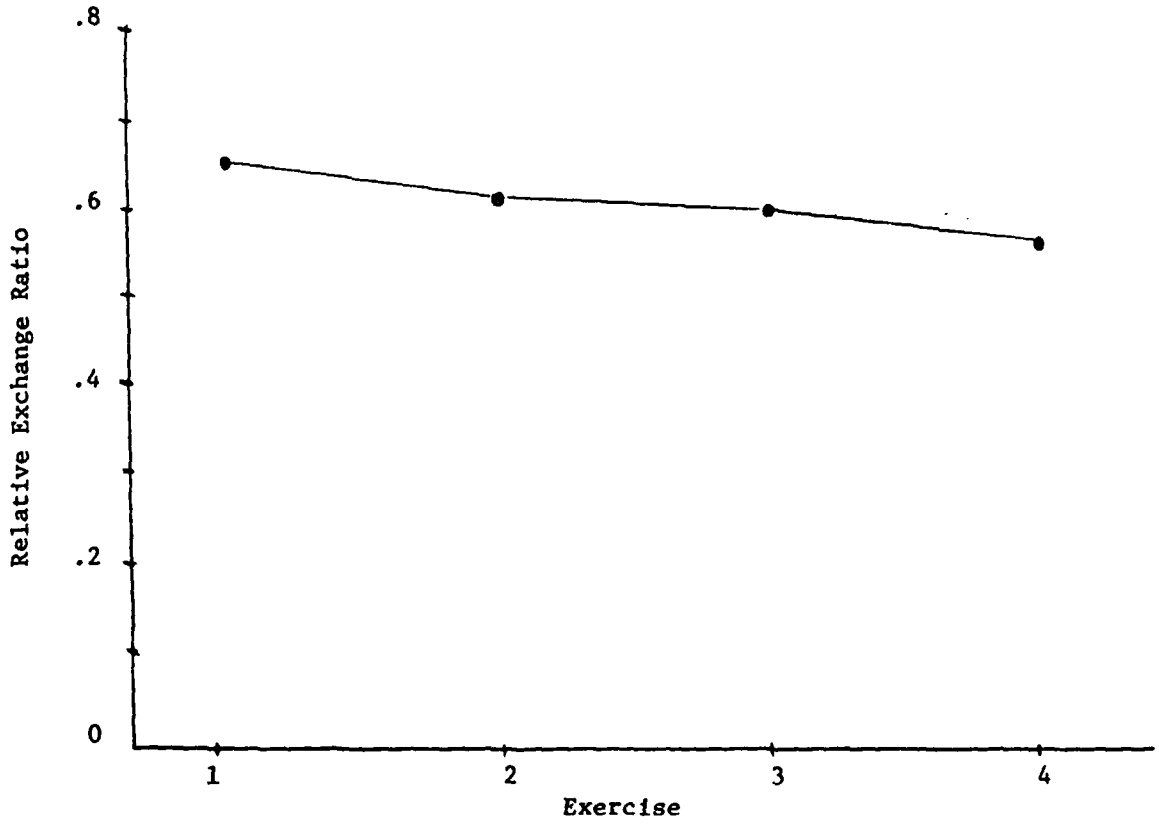

Figure 8. Representative example of simulation outcome measures over successive exercise days. Relative exchange ratio is the percent of OPFOR losses divided by the percent of friendly losses.

Information Flow. Figure 9 shows the effect of exercise day on three communication measures derived from the information flow questionnaire. As explained in the method section, direct reception refers to the recall of information by the command group members of information presented to them by the brigade controllers. Intra-group communication refers to the reception of information by command group members of information that should have been transmitted to them by other command group members. Communication to company commanders refers to the reception by the company commanders of information that should have been presented to them by members of the command group.

All three communication measures increased a small amount from the first administration of the questionnaire to the second (exercise 1 to exercise 3). However, difference t-tests showed that the changes were not statistically significant. The increase might have been larger if the players had received feedback about their communication patterns.

The effects of system characteristics on information flow was not measured, because the information-flow questionnaire was designed to assess communication that takes place primarily during the planning phase of the operations, whereas, any effect of system characteristics on communication would occur primarily during the execution phase. Furthermore, the effects of characteristics were confounded, because the questionnaire was administered only after the second and fourth days.

Fidelity of CATTS Simulation

In no instance were the realism ratings made by the observers (controllers, players, and player-controllers) below the standard for "minimum realism to be of training value in a training exercise." This indicates the fidelity of CATTS in all the areas surveyed, as assessed by these ratings, was at least adequate and sometimes very good.

It was noted that controller ratings of performance were highly correlated with ratings of realism (typically in the mid .80's), further indicating an overall bias in controller ratings. Player-controller correlations between the ratings was also significant, but typically not as great, (usually about r=.60). Finally, correlations for player ratings were even lower, typically in the mid 40's.

GENERAL DISCUSSION

Manipulation of system characteristics had a strong impact on simulation outcomes. These results demonstrate the need for setting levels of these variables so that different combinations of combat ratio, mission weather, communications, and terrain can result in exercises of approximately the same level of battle difficulty. Performance measurement research with subsequent BCGs can, therefore, be facilitated since performance change as a result of CATTS exercises can be attributed to factors other than exercise difficulty when the latter is controlled. Further, it is, in principle possible to match exercise difficulty to the ability levels of individual BCGs in subsequent

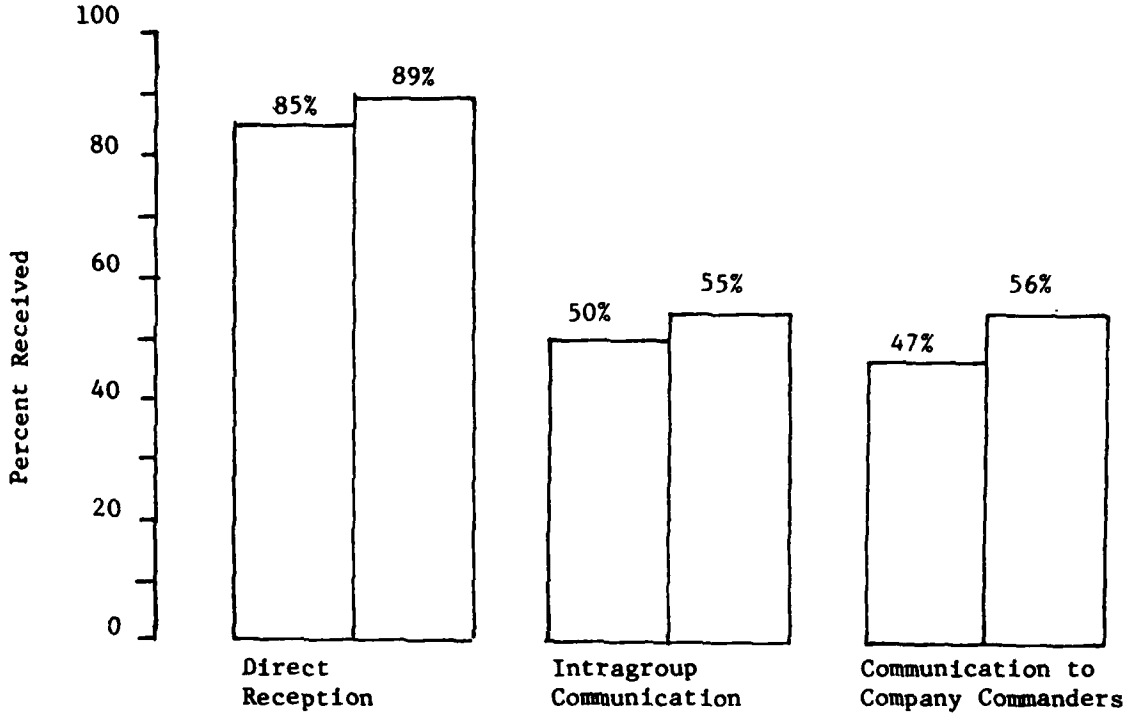

Figure 9. Mean information flow scores for exercise days 2 and 4.

training efforts. In addition, the responsiveness of simulation outcomes to the above system characteristics lends further credibility to the measures as potential indices of BCG performance.

Perception of BCG performance and exercise difficulty were not, however, affected by system and scenario manipulations. It should be pointed out that the experimental design used in this research does not allow for a direct comparison of these measures for player and player-controllers ratings, since raters and experimental conditions were confounded. Controllers did observe all exercises, but their ratings indicate an inability to discriminate among and between items intended to measure BCG performance and exercise difficulty. Each controller was responsible for a different brigade level function which may have resulted in different perspectives among controllers in observing CATTS exercises. This, plus the fact that controllers' duties may have competed for their attention required to observe players, may have contributed to the limited amount of agreement in ratings of BCG performance by controllers.

In terms of performance change across exercise days, the two potential performance measures are inconsistent. While players and controllers perceived a performance increase, simulation outcomes indicated no change. It is possible that players working in the TOC were not strongly affected by events on the simulated battlefield, since information regarding battlefield events were filtered through company commanders. In addition, a high proportion of the performance items were concerned with the planning phases of operations and may not have been affected by factors that primarily affect execution. Also, it is possible that the items rated by controllers, players, and player-controllers were not specific enough (i.e., no clear standards or observable behavioral indicators) to be related directly to battlefield outcomes. Finally, it is possible that manipulations in system characteristics only directly influenced the environment in which BCGs exercised and did not directly influence the C^2 processes engaged in by BCGs. For example, BCGs planned and executed the same procedures regardless of the combat ratio they were confronted with.

In agreement with simulation outcomes, communication did not improve significantly across exercise days. It is acknowledged that the later finding is based on only four BCGs. But, of more importance is the fact that precautions were taken during this research to minimize learning by BCGs across CATTS exercises. That is to say, performance feedback to BCGs in terms of battlefield outcomes, information flow, etc., was kept at a minimum. This was done to reduce the likelihood that battlefield outcomes would be affected by learning and not exclusively by the system characteristics.

Simulation outcomes and information flow did not improve in this relatively controlled experiment. Where exercise difficulty is controlled and steps are taken to encourage learning in CATTS exercises, increases in these measures are more likely to indicate a real training increment as a result of CATTS exposure.

The planned follow-on research intended to demonstrate a C^2 performance increase as a function of CATTS exposure could use a pre-test, training, and post-test format. Pre- and post-test CATTS exercises could be equated for battle difficulty based upon the findings from the simulation outcomes. Performance feedback on battle outcomes and information flow could be provided to the BCGs at the conclusion of the pre-test to enhance subsequent performance on these measures. BCGs would then be exposed to the typical training that is provided at CATTS, and their performance measured on the post-test exercise to assess changes.

It appears that further research is required to develop subjective ratings of BCG performance that are both reliable and valid. Steps should be taken to develop new, more objective measures of C^2 performance, such as a series of naturally occurring battlefield events (probes), that could be used to determine how well BCGs exchange this critical information in a timely fashion during CATTS exercises. The use of performance appraisals by non-involved, outside observers such as TOC monitors to assess C^2 behavior may produce more reliable and valid measures. The results of these measurement techniques should also be provided to BCG's, so that they can modify their performance in appropriate ways. These findings and recommendations will provide the basis for the remainder of the TDS research effort.

REFERENCES

Army Training and Evaluation Program for Mechanized Infantry/tank Task Forces, No. 71.2, Washington, D.C. Headquarters, Department of the Army, 17 June 1977.

Barber, H.F., & Kaplan, I.T. Battalion Command Group Performance in Simulated Combat. ARI Technical Paper 353. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences, March 1979. (AD A070 089)

Barber, H.F., & Solick, R.E. MILES Training and Evaluation Test, USAREUR: Battalion Command Group Training. ARI Research Report 1290. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences, June 1980. (AD A109 192)

Daniel, C. Applications of Statistics to Industrial Experimentation. New York: John Wiley and Sons, 1976.

Kaplan, I.T. Information Flow in Battalion Command Groups. ARI Technical Paper 499. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences, June 1980. (AD A109 469)

Kaplan, I.T., & Barger, H.F. Training Battalion Command Groups in Simulated Combat: Identification and Measurement of Critical Performance. ARI Technical Paper 376, Alexandria VA: U.S. Army Research Institute for the Behavioral and Social Science, June 1979. (AD A075 414)

Stevens, S.S. Psychophysics. New York: John Wiley and Sons, 1973.

Thomas, G.S. Battle Simulation Outcomes as Potential Measures of Battalion Command Group Performance in CATTS Exercises. ARI FLv FU Working Paper 83-1, March 1983.

USACDC, Force Development: The Measurement of Effectiveness, USACDC Pamphlet No. 71-1, 1973.

APPENDIX A The ARTEP Questionnaire With Performance Standards

CONTROLLER'S END OF DAY QUESTIONNAIRE

Name and Rank _____

Position _____

Date _____

Give a number to each item that reflects how many times the item is greater or lesser than the standard. For example, if you think the item is twice the standard, then give it a 200; if you think the item is 1/3 the standard, give it a 33. You may use any number except zero. The standard has a value of 100.

If you did not observe an item during today's exercise, give the item a N/A. Do not agonize over the number for an item. Put down your first impression.

PERFORMANCE

Rate how well this battalion command group performed.

Standard:

100	The minimum acceptable performance in a tactical environment.

I. Gather and analyze required information.

_____ Analyze mission.
_____ Determine what information is available and what additional information is required.
_____ Determine what information sources are available.
_____ Gather all available information and request additional information as needed.

II. Develop a plan based on mission and modify it as required by events.

_____ Determine friendly capabilities and limitations, request additional assets if needed.
_____ Estimate enemy capabilities and likely course of action.
_____ Identify key terrain.
_____ Select battle position/routes to objectives.
_____ Identify critical place.
_____ Develop and compare courses of action.
_____ Individual staff planning: Commo
_____ Individual staff planning: Intel
_____ Individual staff planning: Operations
_____ Individual staff planning: Admin/log
_____ Individual staff planning: Fires
_____ Coordinate with other staff members.

III. Communicate/coordinate.

_____ Issue a warning order.
_____ Disseminate plans and orders.
_____ Disseminate combat information and intelligence.

IV. Implement plan.

_____ Concentrate/shift combat power.
_____ Reinforce terrain.
_____ Provide supplies.
_____ Maintain equipment.
_____ Request additional assets.

V. Supervise combat operations.

_____ Compare battlefield events with current order and concept of operations.
_____ Determine that a new course of action is necessary.
_____ Determine that a change in implementation is necessary.

_____ OVERALL, HOW WELL DID THE COMMAND GROUP PERFORM?

Please elaborate on why you rated any of the above tasks especially high or especially low on the back of this sheet.

Based on your observations of his performance, indicate how well each of the following individuals performed during this exercise. Give a number that reflects how many times better or poorer than the standard the performance was.

Standard:

100	The minimum acceptable performance in a tactical environment.

Position	Performance
Battalion Commander	_____
S1	_____
S2	_____
S3	_____
S4	_____
FSO	_____
ALO	_____
Commander, Team _____	_____
Commander, Team _____	_____
Commander, Team _____	_____
Commander, Team _____	_____

Please elaborate on why you rated any of the above individuals especially high or especially low on the back of this sheet.

DIFFICULTY

Compared to your previous training or tactical experience, rate the difficulty of today's functions.

Standard:

100	Difficulty of functions that is expected in a tactical environment.

	Planning	Execution
Commo	_____	_____
Intel	_____	_____
Operations	_____	_____
Admin/log	_____	_____
Fires	_____	_____
Coordinate with other staff members.	_____	_____

_____ Overall, how difficult was the planning?

_____ Overall, how difficult was the execution?

Please elaborate on why you rated any of the above tasks especially high or low on the back of this sheet.

REALISM

Rate how realistically today's tasks were simulated.

Standard:

100	Minimum realism to be of training value in a training exercise.

	Planning	Execution
Commo	_____	_____
Intel	_____	_____
Operations	_____	_____
Admin/log	_____	_____
Fires	_____	_____
Coordinate with other staff members.	_____	_____

_____ Overall, how realistic was the planning?

_____ Overall, how realistic was the execution?

Please elaborate on why you rated any of the above tasks especially high or especially low on the back of this sheet.

PERFORMANCE STANDARDS

I. Gather and analyze required information.

A. Analyze mission.

Command group identifies specified and any implied tasks or constraints; addresses those tasks during the development of plan and in its own oral warning order/FRAGO/OPORD.

B. Determine what information is available and what additional information is required.

Command group determines information available on area of operations, friendly situation and opposing force situation.

Command group examines information on avenues of approach to objective; type, composition and location of TF and opposing maneuver and fire support units; TF and opposing force units capability of reinforcing by maneuver and fire; location of obstacles and opposing forces air attack and EW capability.

C. Determine what information sources are available.

Command group examines all appropriate resources. All assigned, attached, or DS units as well as higher echelon sources should be considered.

D. Gather all available information and request additional information as needed.

Command group gathers information on opposing force situation, areas of operations, and friendly situation. Records and displays are maintained and additional information requested as necessary from sources identified in Task 1C. As a minimum, collection efforts should focus on determining opposing force intentions and on determining status and situation of TF elements, major adjacent units and brigade reserve and supporting forces.

II. Develop a plan based on mission and modify it as required by events.

A. Determine friendly capabilities and limitations, request additional assets if needed.

Command group analyzes friendly capabilities (in terms of personnel, equipment and supplies, organic/attached/non-organic fires, maintenance and other supporting assets) to support current and contingency plans and requests additional assets if needed.

B. Estimate enemy capabilities and likely courses of action.

Command group, based on an understanding of known opposing force tactics and doctrine, compares that with combat information and intelligence received to determine opposing force intentions.

C. Identify key terrain.

Terrain which facilitates accomplishment of the TF mission is selected for control by occupation or fires. Terrain which, if captured/ controlled by opposing forces, would facilitate accomplishment of the opposing force mission is designated key terrain.

D. Select battle position/routes to objectives.

Command group/staff selects appropriate battle positions/routes to objective which reflects the commander's concept of accomplishment of the mission.

E. Identify critical place.

Command group determines the place on the battlefield where the TF combat power should be concentrated. Comparison made with OPFOR or chief controller's determination of critical place.

F. Develop and compare courses of action.

Based on the command group's analysis of the current situation and mission, courses of action are recommended and compared. The commander selects a course of action and clearly states his concept of the operation.

G. Individual staff planning: Commo.

Communications plan satisfies mission requirements, provides for COMSEC, specifies alternate means of communications, includes the MIJI plan, and can be accomplished with the time and resources available to the TF.

H. Individual staff planning: Intel.

Plan provides for analysis of AO, intel estimates, intel requirements, minimizing TF vulnerability to mass destruction weapons, detecting impending threats to TF security and deceiving opposing force as to TF intentions.

I. Individual staff planning: Operations.

Command group task organizes the task force into company teams. A scheme of maneuver is developed to apply maximum combat power at the critical place while minimizing TF vulnerability. Operational security is addressed in plan.

J. Individual staff planning: Admin/log.

Admin/log plan must complement scheme of maneuver. Priority of support (e.g., replacements, health services, classes of supply) is established. Planning must be flexible to allow changes during execution and facilitate future operations.

K. Individual staff planning: Fires.

Plan provides for preplanned fires, fires against targets of opportunity, suppression, surprise and deception, and air defense coverage while allowing TF elements to maneuver freely.

L. Coordinate with other staff members.

Command group selects control measures which support the scheme of maneuver, facilitate fire and movement by the TF and permit rapid changes as the battle develops. Command group examines components of the plan to identify areas that place limitations on, or require modification in, other components and resolve conflicts.

III. Communicate/coordinate.

A. Issue a warning order.

Upon receipt of a mission, warning order is issued to all necessary subordinate elements. Warning order includes nature of the operation and when and where TF OPORD will be issued. Communications and electronic security measures are rigidly adhered to throughout the TF.

B. Disseminate plans and orders.

Orders are coordinated with appropriate agencies. Orders are issued so as to allow TF elements maximum time for troop-leading procedures. Orders are appropriate, clear, and concise and contain essential information. Changes in plans are communicated orally as a frag order and include changed objectives, control measures, and scheme of maneuver. Communications and electronic security measures are rigidly adhered t- throughout the TF.

C. Disseminate combat information and intelligence.

Combat information should be event oriented rather than in periodic reports and summaries. Only information usable to the recipient should be disseminated. Information should be accurate and disseminated in time for the recipient to act upon it. Communications and electronic security measures are rigidly adhered to throughout the TF.

IV. Implement plan.

A. Concentrate/shift combat power.

Command group assesses the developing situation and issues directives to maximize use of combat power, requesting additional assets if necessary.

B. Reinforce terrain.

Command group tasks TF elements and supporting engineer units to support the scheme of maneuver, reduces TF vulnerabilities, and increases OPFOR vulnerabilities (e.g., minefields and other obstacles, fortifications, etc.).

C. Provide supplies.

Coordinate with supporting supply elements to ensure that supplies (type and number), required to support TF elements and their weapon systems, are available to the TF on a timely basis.

D. Maintain equipment.

Command group determines status of equipment and directs repair/evaluation of non-operational equipment critical to mission accomplishment.

E. Request additional assets.

Command group supervises acquisition, control and expeditious movement of replacements to points where they are needed. Command group requests reinforcements, if required, and supervises transfer of control.

V. Supervise combat operations.

A. Compare battlefield events with current order and concept of operations.

Command group monitors the developing situation, insuring that TF elements and supporting units comply with plans and orders, and that the friendly course of action continues to be appropriate.

B. Determine that a new course of action is necessary.

Command group detects a change in opposing force intentions, a threat to TF security or an inability to support their current course of action with available assets and determines that a change in plan is required.

C. Determine that a change in implementation is necessary.

Command group determines that some aspects of the plan are not being successfully implemented and makes necessary corrections.

APPENDIX B Rationale For Simulation Outcomes

The relative exchange ratio (RER) and surviving maneuver ratio differential (SMFRD) are variations of combat effectiveness measures that were used in previous combat modeling and were found to be useful measures in CATTS exercises. RER is simply the proportion of the percentage of OPFOR losses to the percentage of friendly losses. SMFRD is calculated by subtracting the percent of OPFOR surviving battle from the percent of friendly forces surviving. Both measures are, therefore, rather straight-forward comparisons of losses or surviving strengths of opposing forces.

The C^2ILL ratio is based on the assumption that it is perferable to have a high percentage of forces surviving, while attriting a relatively high proportion of enemy forces. Hence, C^2ILL is basically computed by adding the two components together. The percent of friendly forces surviving component is divided in half because it has been observed that in covering force missions, controller ratings of performance are more responsive to the amount of enemy forces attrited than to the amount of friendly forces surviving. The above weighting factor places a relatively higher emphasis on OPFOR losses.

The change in combat ratio (ΔCR) measure is based on the assumption that it is preferable to end battle with a more beneficial combat ratio than existed prior to battle. The measure, therefore, computes the relative change in combat ratio with respect to initial combat ratio.

In all the above measures, higher positive or lower negative values are preferable from the standpoint of friendly forces. All measures of initial strengths and losses are based on equipment and not personnel. All types of combat equipment were considered in calculations, where tanks, apc's, tows, etc., were combined. Combination was achieved by adding the products of each equipment type and its corresponding combat effectiveness weight. The CATTS battle calculus included combat effectiveness weights for each piece of equipment based on its ability to destroy other types of equipment. For example, an M60A1 tank had a weight of 73, an M113 apc a weight of 19, and a T-62 a weight of 80.

APPENDIX C The Information Flow Questionnaire

POSITION DURING THE EXERCISE _____ DATE _____

Purpose: This questionnaire is designed to provide information to assist in the analysis of lessons learned during CATTS exercises.

Instructions: Answer all questions. Circle the letter that corresponds to what you believe is the correct answer. DO NOT GUESS!! If you did not receive the information asked for, answer unknown. Parts of this questionnaire will cover material to which you did not have access.

1. TF 1-77 passed through which task force in the attack?	a. TF 1-79 b. TF 1-3 c. TF 1-4 d. Unknown
2. ADA support was:	a. 1 Vulcan Plat b. 2 Vulcan Plats c. 1 Chaparral Plat d. Unknown
3. How many air sorties were expected for support of TF 1-77?	a. 2-4 b. 6-8 c. 10-12 d. Unknown
4. ADA alert status at the beginning of the exercise was:	a. Yellow/Free b. Yellow/Tight c. Red/Tight d. Unknown
5. 3rd Bde could expect what level of air support?	a. Air parity b. Limited periods of air superiority c. Enemy air superiority d. Unknown
6. Evacuation of deceased personnel was accomplished _____.	a. With wounded personnel b. By air evac c. Separate from wounded personnel d. Unknown
7. Leaves and passes were _____.	a. Revoked b. Authorized by the company commander c. Authorized by the Bn commander d. Unknown
8. Personnel expected to require hospitalization over _____ hours were sent out of country.	a. 48 hrs b. 72 hrs c. 24 hrs d. Unknown
9. Graves registration personnel required _____ to be sent with all deceased personnel.	a. Nothing b. Personal effects c. Personal effects with individual weapons d. Unknown

10.	A _____ shortage of medical supplies was anticipated.	a. Moderate b. Critical c. Minimal d. Unknown
11.	Before evacuation, POWs were interrogated by the _____.	a. Bn S-2 b. Bde S-2 c. Bn Cdr d. Unknown
12.	Commanders enforced consumption of _____.	a. Hot meals b. Salt tablets c. MCIs d. Unknown
13.	The main attack was directed against the _____.	a. Giddi Pass b. Mitla Pass c. Shu el Hagg d. Unknown
14.	In the 52nd Mech Sector, there were _____ bridgeheads established.	a. Two b. Three c. Four d. Unknown
15.	TF ABE attrited the lead regiments to _____ strength.	a. 20% b. 40% c. 60% d. Unknown
16.	Air Force interdiction of the crossing sites was _____.	a. Successful b. Unsuccessful c. Marginally successful d. Unknown
17.	Enemy's defense today against the 52nd Mech consisted of forces the following size:	a. 2 Motorized rifle divisions b. 2 motorized rifle regiments c. 2 motorized rifle battalions d. Unknown
18.	Enemy activity in the eight hours preceding today's operation consisted of _____.	a. Limited counterattacks b. H&I artillery and reconnaissance c. Major offensives d. Unknown
19.	Attachments should come _____.	a. With their CSS support b. Without their CSS support c. With CSS support as needed d. Unknown
20.	DS contact teams were _____.	a. Assigned to your battalion for operations b. Available upon request c. Not available d. Unknown

21. The basic load for a TOW was _____rds.
   a. 9
   b. 12
   c. 16
   d. Unknown

22. No major end item replacement for at least _____ hours.
   a. 24
   b. 36
   c. 48
   d. Unknown

23. Elements were to maintain a _____ day supply of MCI's.
   a. 2
   b. 3
   c. 5
   d. Unknown

24. Immediate smoke covers _____.
   a. 150-600m for up to 15 min
   b. 150m for 5 min
   c. Other than a. or b.
   d. Unknown

25. The AN/MPQ-4A radar was collocated and reported thru _____.
   a. DIVARTY
   b. 3d Bde FSO
   c. C 2-618 FA
   d. Unknown

26. Arty should be used on targets of _____ or more armored vehicles.
   a. One
   b. Two
   c. Three
   d. Unknown

27. CAS aircraft ordnance loads were weighted in favor of:
   a. Antiarmor
   b. Antipersonnel
   c. Road cratering
   d. Unknown

28. Air observers were available _____.
   a. 160800
   b. 161800
   c. 170300
   d. Unknown

29. Engineer support in the attack consisted of _____.
   a. One squad
   b. One platoon
   c. One company
   d. None

30. The M173 Line Charge was _____.
   a. Not available
   b. In short supply
   c. Available in limited quantities
   d. Unknown

3

31. The engineer element was reinforced with _____.
    a. One dozer
    b. One grader and one CEV
    c. One CEV and one AVLB
    d. Unknown

32. Priority of engineer effort should be_____.
    a. Mobility
    b. Countermobility
    c. Survivability
    d. Unknown

33. All obstacles encountered should be_____.
    a. By-passed if possible
    b. Assault breached
    c. Deliberately breached
    d. Unknown

Number of months in current position with this unit? _____

Number of months in current unit? _____

Previous experience with Battle Simulations?

	Yes	No	Number of Times Played
CATTS			
CAMMS			
BATTLE			
PEGASUS			
CPX			
FTX			
Other (specify)			

Which of the above exercises did you play last as a member of this command group?

_____

When? _____ _____
(month) (year)

APPENDIX D Factor Loadings of Performance and Difficulty Items

Controllers Performance Ratings

FACTOR MATRIX USING ALPHA FACTOR

	Factor I
I. Gather and analyze required information.	
Analyze mission.	.808
Determine what information is available and what additional information is required.	.837
Determine what information sources are available.	.835
Gather all available information and request additional information as needed.	.828
II. Develop a plan based on mission and modify it as required by events.	
Determine friendly capabilities and limitations, request additional assets if needed.	.870
Estimate enemy capabilities and likely course of action.	.791
Identify key terrain.	.739
Select battle position/routes to objectives.	.799
Identify critical place.	.802
Develop and compare courses of action.	.814
Individual staff planning: Commo	.761
Individual staff planning: Intel	.866
Individual staff planning: Operations	.910
Individual staff planning: Admin/log	.761
Individual staff planning: Fires	.903
Coordinate with other staff members.	.893
III. Communicate/coordinate.	
Issue a warning order.	.747
Disseminate plans and orders.	.840
Disseminate combat information and intelligence.	.861
IV. Implement plan.	
Concentrate/shift combat power.	.864
Reinforce terrain.	.822
Provide supplies.	.783
Maintain equipment.	.725
Request additional assets.	.708
V. Supervise combat operations.	
Compare battlefield events with current order and concept of operations.	.861
Determine that a new course of action is necessary.	.836
Determine that a change in implementation is necessary.	.821

Player Performance Ratings

VARIMAX ROTATED FACTOR MATRIX

		Factor 1	Factor 2	Factor 3
I.	Gather and analyze required information.			
	Analyze mission.	.551	.228	.642
	Determine what information is available and what additional information is required.	.420	.339	.748
	Determine what information sources are available.	.408	.397	.690
	Gather all available information and request additional information as needed.	.272	.445	.801
II.	Develop a plan based on mission and modify it as required by events.			
	Determine friendly capabilities and limitations, request additional assets if needed.	.630	.335	.542
	Estimate enemy capabilities and likely course of action.	.606	.207	.560
	Identify key terrain.	.629	.154	.484
	Select battle position/routes to objectives.	.694	.207	.454
	Identify critical place.	.731	.360	.240
	Develop and compare courses of action.	.637	.404	.336
	Individual staff planning: Commo	.538	.415	.347
	Individual staff planning: Intel	.634	.480	.355
	Individual staff planning: Operations	.742	.480	.304
	Individual staff planning: Admin/log	.618	.364	.300
	Individual staff planning: Fires	.456	.613	.514
	Coordinate with other staff members.	.567	.485	.497
III.	Communicate/coordinate.			
	Issue a warning order.	.672	.228	.359
	Disseminate plans and orders.	.644	.457	.317
	Disseminate combat information and intelligence.	.654	.555	.236
IV.	Implement plan.			
	Concentrate/shift combat power.	.576	.633	.256
	Reinforce terrain.	.692	.494	.177
	Provide supplies.	.284	.840	.228
	Maintain equipment.	.272	.846	.121
	Request additional assets.	.232	.761	.242
V.	Supervise combat operations.			
	Compare battlefield events with current order and concept of operations.	.363	.729	.419
	Determine that a new course of action is necessary.	.341	.734	.388
	Determine that a change in implementation is necessary.	.332	.717	.389

Player-Controller Performance Ratings

VARIMAX ROTATED FACTOR MATRIX

	Factor 1	Factor 2
I. Gather and analyze required information.		
Analyze mission.	.566	.614
Determine what information is available and what additional information is required.	.638	.613
Determine what information sources are available.	.789	.410
Gather all available information and request additional information as needed.	.745	.479
II. Develop a plan based on mission and modify it as required by events.		
Determine friendly capabilities and limitations, request additional assets if needed.	.707	.547
Estimate enemy capabilities and likely course of action.	.747	.459
Identify key terrain.	.719	.431
Select battle position/routes to objectives.	.663	.524
Identify critical place.	.576	.626
Develop and compare courses of action.	.625	.607
Individual staff planning: Commo	.468	.582
Individual staff planning: Intel	.724	.320
Individual staff planning: Operations	.672	.593
Individual staff planning: Admin/log	.440	.711
Individual staff planning: Fires	.663	.449
Coordinate with other staff members.	.523	.641
III. Communicate/coordinate.		
Issue a warning order.	.703	.247
Disseminate plans and orders.	.644	.440
Disseminate combat information and intelligence.	.696	.388
IV. Implement plan.		
Concentrate/shift combat power.	.489	.709
Reinforce terrain.	.335	.846
Provide supplies.	.408	.742
Maintain equipment.	.443	.712
Request additional assets.	.614	.611
V. Supervise combat operations.		
Compare battlefield events with current order and concept of operations.	.454	.840
Determine that a new course of action is necessary.	.320	.832
Determine that a change in implementation is necessary.	.471	.709

Controller Difficulty Ratings

VARIMAX ROTATED FACTOR MATRIX

		Factor 1	Factor 2
Planning:	Commo	.876	.334
	Intel	.918	.355
	Operations	.883	.365
	Admin/log	.852	.405
	Fires	.918	.331
	Coordinate with other staff members	.820	.361
Execution:	Commo	.354	.862
	Intel	.551	.812
	Operations	.499	.807
	Admin/log	.299	.911
	Fires	.467	.826
	Coordinate with other staff members	.262	.902
Overall, how difficult was the planning?		.880	.433
Overall, how difficult was the execution?		.371	.916

Player Difficulty Ratings

VARIMAX ROTATED FACTOR MATRIX

		Factor 1	Factor 2
Planning:	Commo	.107	.940
	Intel	.730	.337
	Operations	.782	.370
	Admin/log	.565	.596
	Fires	.617	.613
	Coordinate with other staff members	.591	.490
Execution:	Commo	.272	.451
	Intel	.818	.249
	Operations	.797	.234
	Admin/log	.443	.456
	Fires	.604	.521
	Coordinate with other staff members	.803	.259
Overall, how difficult was the planning?		.697	.514
Overall, how difficult was the execution?		.590	.298

Player-Controller Difficulty Ratings

VARIMAX ROTATED FACTOR MATRIX

		Factor 1	Factor 2
Planning:	Commo	.829	.278
	Intel	.860	.306
	Operations	.868	.304
	Admin/log	.875	.268
	Fires	.837	.386
	Coordinate with other staff members	.896	.306
Execution:	Commo	.390	.559
	Intel	.340	.478
	Operations	.257	.836
	Admin/log	.169	.826
	Fires	.393	.702
	Coordinate with other staff members	.498	.678
Overall, how difficult was the planning?		.779	.464
Overall, how difficult was the execution?		.259	.852

APPENDIX E Source Tables for F^2 Analysis of Simulation Outcomes

Source Tables for F^2 Analysis of Simulation Outcomes

RER

Source	Sum of Squares	df	Mean Squares	F	P
Total	4973.84	15			
Terrain	.80	1	.80	-	
Mission	1596.20	1	1596.20	18.43	<.005
Weather	491.84	1	491.84	5.68	<.05
Commo	116.15	1	116.15	1.34	
Combat Ratio	1404.56	1	1404.56	16.22	<.005
Mission X Weather	86.54	1	86.54	< 1	
Mission X Terrain	338.84	1	338.84	3.91	<.10
Mission X Commo	332.61	1	332.61	3.84	<.10
Mission X C.R.	25.16	1	25.16	< 1	
Residual	581.13	6	96.86		

APPENDIX E (Continued)

SMFRD

Source	Sum of Squares	df	Mean Squares	F	P
Total	183.98	15			
Terrain	1.19	1	1.19	-	
Mission	8.74	1	8.74	2.10	
Weather	22.20	1	22.20	5.33	< .05
Commo	10.61	1	10.61	2.55	
Combat Ratio	93.94	1	93.94	22.55	< .005
Mission X Weather	3.54	1	3.54	< 1	
Mission X Terrain	9.80	1	9.80	2.35	
Mission X Commo	1.48	1	1.48	< 1	
Mission X C.R.	4.10	1	4.10	< 1	
Residual	28.37	6	4.73		

APPENDIX E (Continued)

ΔCR

Source	Sum of Squares	df	Mean Squares	F	P
Total	13947.30	15			
Terrain	146.53	1	146.53	-	
Mission	2522.05	1	2522.05	9.21	< .025
Weather	1165.20	1	1165.20	4.26	< .10
Commo	441.42	1	441.42	1.61	
Combat Ratio	6143.42	1	6143.42	22.43	< .005
Mission X Weather	1.97	1	1.97	< 1	
Mission X Terrain	84.36	1	84.36	< 1	
Mission X Commo	5.48	1	5.48	< 1	
Mission X C.R.	1064.06	1	1064.06	3.89	< .10
Residual	2372.81	6	395.47		

APPENDIX E (Continued)

C^2ILL

Source	Sum of Squares	df	Mean Squares	F	P
Total	74.46	15			
Terrain	.01	1	.01	-	
Mission	9.31	1	9.31	6.33	< .05
Weather	8.93	1	8.93	6.07	< .05
Commo	5.63	1	5.63	3.82	< .10
Combat Ratio	32.04	1	32.04	21.78	< .005
Mission X Weather	.70	1	.70	< 1	
Mission X Terrain	4.69	1	4.69	3.19	
Mission X Commo	2.03	1	2.03	1.38	
Mission X C.R.	1.52	1	1.52	1.03	
Residual	9.60	6	1.60		

APPENDIX F

Mean Performance, Planning Difficulty, and Execution

Difficulty Ratings by Controllers as a Function of System Characteristics

	Combat Ratio		Mission		Weather		Terrain		Commo	
	Good	Bad	Attack	CF	Good	Bad	Fulda	Sinai	Good	Bad
Performance	120	117	121	116	113	124	118	119	120	117
Planning Difficulty	108	116	112	112	112	111	115	109	109	115
Execution Difficulty	108	107	104	112	118	99	106	111	108	107

APPENDIX G

Mean Performance and Difficulty Ratings Across Exercise Days

	Performance				Difficulty				
Exercise	1	2	3	4	Exercise	1	2	3	4
Controllers					Controllers				
Factor I ***	87	107	122	130	Factor I ***	103	110	120	124
					Factor II *	99	111	120	126
Players					Players				
Factor I **	119	136	149	147	Factor I	113	122	121	124
Factor II ***	119	136	153	149	Factor II	97	108	111	109
Factor III *	112	133	150	146					
Player-Controllers					Player-Controllers				
Factor I	114	147	156	151	Factor I *	101	124	137	136
Factor II	116	145	153	150	Factor II	116	126	131	132

* $p<.05$
** $p<.005$
*** $p<.001$

APPENDIX H

Mean Simulation Outcome Scores for First and Second Attack and Covering Force Missions

C^2ILL

	Attack	Covering Force
1st	1.56	1.94
2nd	1.72	1.73

ΔCR

	Attack	Covering Force
1st	-1.75	-2.19
2nd	-1.12	-2.70

RER

	Attack	Covering Force
1st	.76	2.00
2nd	1.12	1.47

SMFRD

	Attack	Covering Force
1st	-1.11	-1.02
2nd	- .78	-1.25